Essential genes and gene products for identifying, developing and optimizing immunological and pharmacological active ingredients for the treatment of microbial infections

Description

The present invention relates to a method for identifying and characterizing essential genes of pathogenic microorganisms, to the use thereof for finding novel immunological and pharmacological active ingredients for the prophylaxis, therapy and diagnosis of bacterial infections, and to the further development and optimization of these active ingredients. The invention includes the corresponding nucleic acids which code for the essential gene products, and the polypeptides encoded thereby. The invention additionally relates to vectors which comprise the nucleic acids of the invention, to cells transformed with these vectors and to antibodies specific for the polypeptides. These nucleic acids and polypeptides can be employed for the diagnosis, prevention and treatment of microbial infections; in particular, they can be used for developing antibodies, vaccines and inhibitors.

The complete molecular elucidation of the human genome and of that of clinically relevant pathogens is opening up new ways of developing agents for the therapy and prophylaxis of human diseases. Thus, deciphering of the human genome is announced to take place in the next few years. The number of pathogenic microbes whose molecular characterization is complete is continually increasing. The stated aim is now to identify from the comprehensive data material those genes whose products are suitable as a potential target for an active ingredient and thus can be used for developing a specific active ingredient. This potential cannot be inferred from the primary structure of a gene but must be determined by experiment.

If the complete genomic sequence of an organism is available, the problem arises of making the enormous amount of data available for further biological analyses. The first step is to identify all the genes located on the genome. This usually takes place with the aid of computer-assisted search programs which are able to predict potential genes with a degree of certainty. It is possible in this way to produce gene maps which, however, are still associated with great inaccuracy. If the genes identified by the search program cannot be assigned to any known gene, it is necessary for the function of these hypothetical genes to be demonstrated by physical detection of the gene products in the original cell.

Another strategy which is likewise based on the use of special search programs aims at the identification of possible gene families which are linked to specific biological properties which in turn, come under consideration, on the basis of further assumptions, as target of active ingredients. The search criteria aim at characteristic structural features which have usually been derived from genes which are already known. The result of such a search may, depending on the closeness of the preconditions to the actual state, provide a higher rate of hits. However, the inaccuracy of this method is usually relatively high, and the actual biological property of the gene or its gene product must always be confirmed by experiment.

A further strategy detects the expression products of a cell, by which means it is possible to identify the genes active at the particular stage of development. Comparison of different stages of development with one another makes it possible to deduce the interaction of the genes and, in some cases, the biological function of unknown genes can be partly deciphered.

If corresponding comparison studies are carried out with cells having a pathological appearance, it is in fact possible even to identify disease-causing genes and to employ them as potential targets for active ingredients in developing active ingredients.

All the described methods are used in particular to identify previously unknown genes and assign a biological function to them with the aid of computer-assisted comparisons of data. However, none of the known methods provides an unambiguous assessment of a gene or a gene product in relation to its potential as target of active ingredients and thus possibly being used for developing active ingredients.

Some of the most important prerequisites for a pathogenic organism to survive in a host and to multiply are, on the one hand, the ability to avoid the host's immune systems and, on the other, the ability to adapt to a very specific habitat or niche. The factors and proteins necessary for this are thus usually essential for the pathogenic microbe.

It would be a great advantage to identify these essential genes of microorganisms in order in this way to acquire the possibility of producing therapeutic, preventive and/or diagnostic agents, for example antibodies, vaccines or inhibitors of the corresponding polypeptides.

A pathogen of particular medical interest is *Helicobacter pylori*. This microbe is a Gram-negative, spiral-shaped bacterium with a high pathogenic potential which has in recent years increasingly developed resistance to a number of therapeutically relevant antibiotics and is thus of great clinical importance. This is distinguished by extremely high motility because of its flagella and the unusual ability to survive in the strongly acidic medium (down to pH 1.5) of the stomach (Goodwin et al., 1989).

Although the occurrence of spiral-shaped bacteria in the human gastric mucosa has been known for a long time, it is only since the successful isolation and cultivation of this bacterium (Warren and Marshall 1983; Marshall et al., 1984) from the gastric mucosa of a patient with a stomach ulcer (ulcus ventriculi) that these have been known to be pathogenic microbes. *H. pylori* infection is one of the

commonest chronic bacterial infections of humans. It occurs worldwide, with about 50% of the population being infected with this bacterium.

An infection inevitably leads to the development of bacterial gastritis (type B gastritis) in humans. It is also assumed that *H. pylori* is also involved in the causation of stomach and duodenal ulcers (ulcus ventriculi and ulcus duodeni) and some forms of gastric carcinoma (adenocarcinoma) (Lee et al., 1993; Solnick and Tompkins, 1993). In two studies in 1991, a statistically significant correlation was shown between *H. pylori* infection and the occurrence of gastric carcinoma (intestinal type), and both studies concluded that about 60% of all gastric carcinomas which occur are probably attributable to *H. pylori* infection (Parsonnet et al., 1991; Nomura et al., 1991). The less commonly occurring MALT (mucosa associated lymphoid tissue) lymphomas of the stomach, which are regarded as precursors of B-cell tumors of the immune system, are also presumed to be a consequence of *H. pylori* infection. One consequence of chronic infection with *H. pylori* is atrophic gastritis, a degeneration of the mucus-, acid- or pepsin-producing cells of the gastric epithelium, which must be regarded as a precancerous lesion.

After oral intake, the bacteria initially reach the extremely acidic lumen of the stomach (pH 1 to 2). There is by the production of the enzyme urease, which leads to cleavage of the urea which is present and thus to local neutralization of the acidic pH in the stomach, which makes it possible for the bacteria to survive. The microbes then move by means of chemotactic orientation and flagella-dependent motility into the bicarbonate-buffered mucus layer of the antrum region of the stomach, their real natural habitat. There they are in the unique ecological niche which, because of the acid barrier, is accessible to only a few competing species of bacteria. The bacteria presumably orient themselves on the pH gradient between lumen (pH 1-2) and epitheal cell surface (pH 6-7) in order to reach the epithelium. Their spiral shape, their motility in the viscous mucus, the production of mucus-modifying enzymes and finally a microaerophilic lifestyle mean that these

microbes are optimally adapted to the living conditions in this habitat. They usually stay in the deep crypts of the antrum region, where they are protected from external influences such as, for example, acid, pepsin, but also from medicaments for eradicating them, such as, for example, antibiotics. Part of the bacterial population (about 20%) is closely associated with epitheal cells, especially with mucus-producing cells. Provided there is gastric metaplasia, that is to say acid-induced development of gastric epithelium in the duodenum, there is also colonization of metaplastic areas in the duodenum, creating the preconditions for the development of the duodenal ulcer (ulcus duodeni). Their adherence ability presumably prevents complete excretion of the *Helicobacter* with the shed mucus, so that the bacteria may persist for years, decades or even life-long (chronic infection).

Before the existence and the significance of *H. pylori* for ulcers were known, the latter were treated by so-called antacids, or H_2 receptor antagonists. These are substances which inhibit the acid secretion by the gastric parietal cell. Although there is usually healing of ulcers under the influence of these pharmaceuticals, because one of the causes of these ulcers, mainly the *H. pylori* infection, is not eliminated thereby, in most cases the ulceration reappears after a short time (recurrence).

Another therapy frequently used for ulcerations is bismuth treatment. Various bismuth salts (CBS, BSS) have a bactericidal effect on *H. pylori*. However, a significant disadvantage of this type of therapy is that total eradication of the microbe is achieved in only a very low percentage of cases (8 to 32%). As with treatment with antacids, there is only temporary suppression of the microbe and, after discontinuation of the treatment, in most cases there is recrudescence of the infection. A further disadvantage of bismuth treatment is that prolonged therapy with high doses leads to accumulation of this substance in the liver, kidney and nervous system and has considerable neurological side effects (Malfertheiner, 1994).

Since the realization that peptic ulcers are infectious diseases, antibiotics are now also employed for treatment. Monotherapy with various antibiotics (amoxicillin, nitrofuran, furazolidine, erythromycin and the like) has, however, proved to be unsatisfactory because even here complete eradication of the microbes occurs in only 0 to 15% of cells. The most successful treatment to date is currently achieved by combination of an acid blocker (ompeprazole) with an antibiotic (amoxicillin), which may lead to eradication rates of up to 80% (Malfertheiner, 1994). However, in the long term, antibiotic treatment for eliminating *H. pylori* is unpromising because rapid development of resistance by the bacteria to antibiotics must be expected because of the incomplete eradication of the microbe.

The increasing occurrence of antibiotic resistances and the limited treatment options, which usually have considerable unwanted side effects, makes it necessary to find new types of therapy and, in this connection, in particular to identify novel active ingredients, especially vaccines, which can be used both for the prophylactic and therapeutic treatment of *Helicobacter* infections. The dosage form is also of particular interest because the active ingredient must be effective in the stomach, that is to say in an extremely acidic environment. Compounds with proton blockers given, for example, before the administration of the prophylactic or therapeutic active ingredient may be of great benefit in this connection.

The molecular basis for persistent chronic *Helicobacter* infections has not as yet been explained. It has been possible to show that the urease, motility and adherence factors are essential properties of the bacterium before being able to colonize the gastric mucosa. Although the host organism is unable under normal conditions to deal with an *H. pylori* infection, it has been shown in an animal model that urease, an essential *H. pylori* virulence factor, has a high potential as a vaccine (US patent application US-SN-07/970,006 "Urease-based Vaccine Against *Helicobacter* Infection).

However, the components responsible for the pathogen being able to avoid the host's immune system are as yet unknown.

Pathogenic organisms have generally developed a large number of strategies for being able to persist in the host untroubled by the immune system over a prolonged period (Haas and Göbel, 1992; Finlay and Falkow, 1997). One mechanism for surviving in a life-threatening environment is to develop into a resistant form.

In the case of *H. pylori*, coccoid forms have frequently been described in the literature as potential resistant forms, but their clinical significance is controversial. Coccoid forms might be very important for *ex vivo* survival. Concerning *in vivo* survival, it has been shown that coccoid forms are preferentially induced by an unfavorable environment such as, for example, a high O_2 partial pressure or sublethal doses of antibiotics (bismuth subcitrate, erythromycin, amoxicillin, metronidazole) (Donelli *et al.*, 1998; Bode *et al.*, 1993; Sorberg *et al.*, 1996; Berry *et al.*, 1995).

Some researchers assume that these coccoid bacteria are viable but non-culturable (VNC). Eaton and coworkers achieved successful infection of minipigs with vegetative (spiral) *H. pylori*, whereas coccoid forms showed no infection in this model (Eaton *et al.*, 1995). Direct detection of coccoid forms in the human stomach was achieved by Chan et al. using sections of stomach tissue from biopsy material. In 82.8% (53/64) of the investigated biopsy samples, the authors were able to detect coccoid forms with *H. pylori* (Chan *et al*, 1994). Cao et al. used a monoclonal antibody for specific detection of coccoid *H. pylori* im in a tissue section. Once again, *H. pylori* coccoid forms were detected, besides the vegetative forms, in 100% of antrum biopsies (9/9) (Cao *et al.*, 1997).

The binding to epithelial cells and the capability for signal transduction (IL-8 induction, rearrangement of the cytoskeleton, binding of plasminogen, lactoferrin

and vitronectin to the bacterial surface) appears to be retained comparably in coccoid forms and vegetative forms (Khin *et al.*, 1996; Segal *et al.*, 1996).

The abovementioned experiments indicate that coccoid forms are important for the survivability of *Helicobacter* in an unfavorable environment. The identification of genes connected with the development of this form and reactivation into the vital form is therefore of the greatest interest for developing novel active ingredients.

Besides *Helicobacter pylori*, it is also possible for other *Helicobacter* species to colonize the human stomach, such as, for example, *H. heilmannii and H. felis.* It has been possible to show in this connection that *H. heilmannii* may also be associated with pathological ulcers. The causative transmission probably takes place from domestic animals to humans. To date, *H. pylori*, which occurs often in humans, has been suspected of being involved in the development of gastric cancer. There are now clinical data which cast doubt on this association. These doubts are supported in particular by recent data on *Helicobacter heilmannii*, which attribute the latter with a greater carcinogenic potential and emphasize its importance in the development of gastric MALT lymphoma (Regimbeau *et al.*, 1988).

Considering what has previously been said in summary makes it clear that there is a need for new types of therapy for controlling bacterial pathogens, in particular for vaccines and inhibitors of essential genes and their expression products. The increasing development of resistance to a large number of proven medicaments requires a continuous supply of novel active ingredients. This increasing need for novel active ingredients can be met only if new active ingredient targets are identified and used to develop novel active ingredients. Essential genes represent an ideal target for the development of active ingredients because they are necessary for the survival of the pathogen.

Therefore one aim of the invention is to identify essential genes of *Helicobacter*, in particular of *H. pylori* or *heilmannii* and of possible resistant forms of *Helicobacter* for the development and optimization of novel therapeutic, preventive and/or diagnostic agents such as, for example, vaccines and pharmacological active ingredients. The priority is to find essential microbial genes, it also being possible to identify homologous proteins of different pathogenic microbes. It would then be possible, as with classical antibiotics, to eliminate a plurality of pathogenic microbes simultaneously by use of one active ingredient. The priority with *Helicobacter* in particular is for genes which fulfill vital functions in the infection process, and genes involved in the development and reactivation of coccoid forms. Of particular interest in this connection are essential genes which code for secreted gene products because the latter can, because of their exposed location, be reached especially well by immunological and pharmacological active ingredients and are therefore good candidates for developing active ingredients. Also of interest are essential genes which code for gene products which are involved in the development and maintenance of resistant forms. A further objective to find essential microbial genes, it also being possible to identify homologous proteins of different pathogenic microbes. It would then be possible, as with classical antibodies, to eliminate a plurality of pathogenic microbes simultaneously by use of one active ingredient.

This object is achieved by a method for providing agents for the detection, therapy or/and the prevention of microbial infections, which comprises the following steps:

(A) identification of essential genes and the corresponding polypeptides by producing gene-deficient microorganisms by conditional antisense inhibition (CAI) or/and subtractive recombination mutagenesis (SRM) and determining the viability and survivability of the gene-deficient microorganisms in an assay system.

(B) identification of specific active ingredients which are directed against the essential polypeptides and bring about inactivation of the microorganisms or used microorganisms.

(C) testing of the identified active ingredients for their usability as components of diagnostic, preventive or/and therapeutic compositions,

(D) formulation of the useful active ingredients as diagnostic, preventive or/and therapeutic compositions.

The method described herein relates to the identification of essential genes and their use for developing novel active ingredients.

A further aspect of the present invention is thus also a method for identifying essential microbial genes, which comprises the following steps:

(i) production of gene-deficient microorganisms,

(ii) determination of the viability or/and survivability of the gene-deficient microorganisms from (i),

(iii) identification of a protein-encoding section of a microbial DNA sequence in which the gene-deficient microorganisms are deficient and

(iv) characterization of those DNA sections which are essential for survivability.

CAI is the abbreviation for conditional antisense inhibition. This is a method which is described in detail hereinafter.

SRM stands for subtractive recombination mutagenesis and is likewise described below.

The term "gene-deficient" as used herein means that the deficient organism is unable to produce one or more of its gene products or utilize the function thereof. The production of the corresponding gene product can, on the one hand, be prevented by mutagenesis of the corresponding gene, or inhibition may take place during expression, for example through antisense nucleic acids. Mutagenesis can be employed to mutate a gene in the genome of the microorganism or to introduce a mutated gene into the microorganism, it also being possible to make use of homologous recombination.

A protein-encoding section of a nucleic acid sequence is, for example, a gene or a part of a gene which permits expression of a polypeptide.

The term "essential gene" means a gene which codes for a gene product without which an organism is unable to survive or is capable of only limited survival. Essential genes can be divided into two classes: obligately essential and facultatively essential genes. An obligately essential gene codes for a protein which is indispensable for the survival or reproduction of an organism under all circumstances. By contrast, a facultatively essential gene codes for a protein which is necessary for the survival or reproduction of the organism only under certain conditions, such as, for example, the ability of the organism to survive inside cultivated mammalian cells or in the animal. In both cases, the survival or the reproduction of the organism is greatly impaired or prevented by inactivation of a gene essential for it or inhibition of a gene product essential for it. If a bacterium is no longer able to survive or its reproduction is impaired after inactivation of a particular gene, this can be regarded as initial evidence that essential properties are mediated by this gene. However, the validity of such findings must be supported by accompanying control experiments, for example it ought to be possible to abolish such a lethal mutation in a second step by an appropriate complementation of the gene or gene product. Obligately essential genes are accordingly those whose nonexpression or absence, for example through mutagenesis or deletion, leads to the organism being non-viable either in the natural environment or on a complete medium which is ideally suited to the needs of the microorganism. If a microorganism is deficient in a facultatively essential gene, it is usually still able to grow on such a complete medium defined according to the organism, but is no longer able to survive in the natural environment, that is to say in its natural host or cells or tissue cultures of its natural host.

Identification of ess ntial gen s

It is possible by the novel method to identify essential genes, irrespective of their specific function, of microorganisms. This method is preferably employed for identifying essential genes from *Helicobacter* and related microorganisms.

In a first step, the complete genome of a bacterial pathogen is screened for essential genes by a molecular genetic approach. This step requires no knowledge at all about the primary structure of the genome or individual genes, but takes place exclusively on the basis of biological criteria. If a gene is identified as essential determinant, its identity is established. Recourse may be had in this connection to the established raw sequence data of the genomic sequencings. It is possible on the basis of the established gene sequence to establish, for example, isogenic variants or whether the established gene is present in an operon in which other essential genes are possibly present.

In a second step, the identified genes are transferred into special genetic systems which are used to provide the genes or their gene products for direct active ingredient screening, and/or the genes or gene products are used for further optimization of previously identified active ingredients. The essential advantage of the overall method is based on the gene and active ingredient screenings being carried out in rapid sequence in informative biological systems so that the potential targets for active ingredients in the complete set of genes of a pathogenic microorganism can in a relatively short time be identified, produced and employed directly for active ingredient screening or optimization.

If a microorganism whose genome has already been sequenced is investigated, a gene or gene section can be identified by means of database analyses, assigning a reading frame to a sequence section. However, it is possible and preferred, irrespective of the presence of a complete genome generation, to subject any desired gene library to a preselection. In this case it is possible and

preferred for the preselection to be carried out for genes which code for polypeptides with a particular function, for example with the aid of homology analyses. The preselection can also be carried out for genes which are expressed only in particular stages of development.

Within the first step it is possible to achieve by selection steps a great reduction in the genetic material to be investigated. For example, by an enrichment step for genes which code for exported or secreted gene products. In this specific method, the DNA sections of a gene library of a pathogen are mutagenized, which can take place, for example, by cloning such a DNA section into a plasmid, transformation into a preferably heterologous host organism and subsequent mutagenesis. The expression product resulting therefrom can then be detected. The mutagenesis can take place, for example, by insertion of a marker sequence which, on expression of the mutagenized sequence in a host organism, leads to a fusion polypeptide for which selection is possible. Insertion of the marker sequence is not restricted to transposon insertion but can also take place in another way, for example by homologous recombination or infection and recombination with the aid of bacteriophages.

The marker sequence used for the purpose of the present invention is generally a gene which codes for a gene product which permits selection for those host organisms which express this sequence. These marker sequences are generally resistance genes which confer resistance to particular antibiotics, or which permit the host organism to survive and reproduce in a selection medium. The genetic marker preferably has no expression signals of its own but depends directly on an upstream promoter such as, for example, the transcription promoter on the promoter segment or a promoter which is located on the cloned heterologous DNA fragment to be identified. Enzymes can also be employed as genetic markers as alternatives to the antibiotic resistance marker sequences. In these cases, successful insertion is indicated by a particular biochemical reaction such as, for

example, a color reaction which allows manual isolation of the corresponding bacterial clone.

If the marker sequence is expressed as fusion protein with the expression product of the inserted DNA fragment, and a selection is carried out as described above, it is possible to isolate DNA material from the selected bacterial clones and to determine the DNA sequence which codes for the fusion product by known methods. This allows a reading frame to be assigned to the DNA fragment to be identified. It is then possible to carry out comparison studies with generally available DNA sequence databases in order to elucidate the identity of the identified gene and, where appropriate, obtain information about a biological function.

A targeted reduction in the sample volume can be achieved by technical and other supplements to the two methods, CAI and SRM, described below. These likewise comprise preceding selection methods aimed at particular groups of genes, for example the use of subtractive gene libraries of pathogenic and apathogenic representatives. Pathogenicity-mediating gene regions are enriched in this way. Subtraction methods of this type can also be used in order to identify genes specific for particular organisms, for example by comparison and subtraction of the genomes of *H. pylori* and *H. heilmannii*.

In further methods it is possible, for example, to identify gene groups which are expressed only in a particular development step. An example which should be emphasized is the array method in which the individual gene probes of the pathogen are applied in the form of a grid to a support. The individual application points are known so that when there is a positive hybridization reaction with the development-specific transcription products or cDNAs or subtractive cDNAs or fragments thereof it is possible for the respective genes to be identified and subsequently cloned. Other methods which detect developmental specific gene groups are comparative proteomic and differential display analyses.

In order to find whether the identified gene sequences are essential genes, microorganisms deficient in the sequences corresponding to the identified gene sequences are produced. The deficient microorganisms are then tested on various growth media or cell cultures or in an animal model or in the natural host, and the deficient genes can then be assigned, depending on the growth ability, to a category of nonessential, obligately essential or facultatively essential genes.

The significance of genes which control the development from the vital to the resistant form and vice versa has already been pointed out. It is therefore particularly preferred to carry out a preselection for such genes. It is subsequently possible to use methods such as CAI or SRM, and the gene-deficient microorganisms are then investigated on particular nutrient media which induce transition from one form into the other. In the case of *Helicobacter,* the Schivo medium which makes reactivation of the coccoid form into the vital spiral form possible is particularly preferred.

Deficient microorganisms can be generated in several ways.

A number of molecular genetic methods are available for mutagenizing the genome of a bacterial pathogen in such a way that a mutant of each gene is available. The commonest mutagenesis method is based on inactivation of genes, for example by random insertion into the genome of transposons which are selected via appropriate markers. Numerous variations of this method which can be applied to different organisms exist. (Joyce and Grindley, 1984; Akerley, *et al.,* 1998). It is also possible with the aid of the inserted transposons to locate the mutagenized gene accurately in the genome.

Once a genetic mutant with a detectable biological effect has been generated, for example a reduced growth of the cells in a particular environment, it is necessary in a second step to demonstrate unambiguous coupling of the gene or of the gene product with this property. This usually takes place by complementation

experiments. In this case, the original gene is introduced into the organism with the specific gene mutants, and is expressed. If it is possible in this way to regenerate the original property of the organism, the necessary proof has been provided. However, this method cannot be applied for the characterization of lethal mutants, that is to say for mutants of obligately essential genes. A way round this is to use conditional mutations for the complementation. For example, it is possible by chemical mutagenesis of the investigated gene to generate temperature-sensitive mutants which convert the gene product at the normally optimal growth temperature into an inactive state, and at lower temperatures, yield a biologically active gene product (Das, *et al.,* 1976; Harris, *et al.,* 1992; Hou, *et al,* 1994; Polissi and Georgopoulos, 1996). In another approach which is practised, the wild-type complementations are controlled by exogenous substances, so-called inducers. These inducers switch on expression of the complementing gene which is introduced on an episome into the gene-specific mutant and, after induction, replaces the missing gene product (Murphy, *et al.,* 1995-. Chow and Berg, 1988; Arigoni, *et al.,* 1998).

Said methods are very time-consuming and are employed only for investigating individual genes or limited genomic sections. Methods which make thorough characterization of the complete genome of a selected pathogen possible by the scheme described have not been disclosed to date.

The novel genetic methods which are described hereinafter, the conditional antisense inhibition (CAI) and the subtractive recombination mutagenesis (SRM), comply with these requirements. Both methods can be employed for identifying essential genes, with the CAI method being particularly suitable for identifying obligately essential genes and the SRM method for facultatively essential genes.

The CAI method is based on the conditional inhibition of the translation of one or more genes located on a clone genome fragment (which then serves as template) and are propagated via a plasmid in the microbe to be investigated. Compared

with conventional methods, the genomic structure of the microbe to be investigated remains unchanged, that is to say in the original state. In the microbe to be investigated, the inhibition of translation is induced by the conditionally inducible synthesis of specific antisense RNA (asRNA) which encompasses the complete cloned genome fragment, including the genes located on the genome fragment. The antisense nucleic acid sequences can then be synthesized in large quantities in the microorganism and bind to the original mRNA, whereupon this mRNA can no longer be translated and is thus removed from the expression apparatus. The consequence is that either no gene product or only small quantities thereof are formed. Synthesis of the asRNA is subject to the control of a promoter (as promoter) whose activity is controlled conditionally, by defined external signals. This conditional inhibition of the expression of a gene or operon thus takes place via regulation of the synthesis of the asRNA by the inducible as promoter. To demonstrate that a gene or operon is, as in the present case, essential for survival and reproduction of the organism under particular conditions, the survival and reproduction rate of a clone in which the synthesis of asRNA induced is compared with its survival/reproduction rate when the asRNA synthesis is not induced. If the survival/reproduction rate of the clone is reduced on induction of asRNA synthesis, then the inhibited gene or operon is an (obligately or facultatively) essential gene. These growth analyses can be carried out automatically, so that a very large number of genes can be investigated within a short time. The plasmid is isolated from the clones, and the DNA sequence of the cloned genome fragment which serves as template for the asRNA synthesis is determined and subsequently the structure of the essential gene is established.

A plasmid vector suitable for the CAI method is depicted in figure1. It contains a genomic or subgenomic DNA fragment from the microorganism to be investigated under the control of an inducible promoter (P_i) and other usual expression signals, and an mRNA-stabilizing element, so that the DNA fragment can be expressed in the form of antisense RNA (asRNA) and has a long biological activity. An example of a suitable promoter is the Tet promoter. In a particularly preferred embodiment,

the CAI vector additionally encodes a gene for a regulatory protein which regulates the promoter, in this case for example the Tet repressor, which can be controlled by an exogenous or extracellular signal such as, for example, tetracycline. The CAI vector of the particular embodiment in figure1 additionally contains one or more selectable marker genes and two origins of replication (ori), one for the microorganism to be investigated (here referred to as pathogen) and another one for a conventional cloning host, for example *E. coli.* Antisense libraries can be constructed from whole microbial genomes with the aid of such CAI vectors.

Figure 2 shows a diagrammatic representation of a preferred CAI method. From a CAI plasmid which contains small fragments of a genomic library of the microorganism to be investigated, asRNA is synthesized, starting from an inducible promoter (P_i), under the control of an extracellular signal (see fig. 1). The asRNA hybridizes sequence-specifically with the mRNA of the gene corresponding to the cloned DNA fragment on the CAI plasmid. Formation of the asRNA-mRNA hybrid reduces or prevents translation of this mRNA. As a consequence there is formation of a deficient microorganism which is unable to form the relevant gene product. If the product is that of an essential gene whose formation is inhibited (A), the viability of the corresponding clone is restricted or impeded. The viability of the microorganism is subsequently determined on the basis of its living or survival or reproduction rate in a defined biological system. If there has been no induction of asRNA synthesis (B), or if the CAI plasmid contains the fragment of a nonessential gene (C), the clone of the microorganism has normal viability and capability of reproduction.

In particularly preferred embodiments of the CAI method, complete antisense RNA plasmid libraries from genomic fragments of the microorganism to be investigated are analyzed (see figure 3). A genomic library with CAI plasmids (see fig. 1) is transferred, under noninducing conditions, into the homologous microorganism to be investigated, and the plasmid-bearing clones are selected via a plasmid-encoded marker. The viability of the individual clones, each of which contain a

particular CAI plasmid from the gene library, is subsequently investigated on the basis of their reproduction rate under induced and noninduced conditions (+ and - in the figure), based on the asRNA synthesis, in direct comparison. Translation of at least one essential gene is prevented in clones which multiply scarcely at all or only slowly under asRNA-inducing conditions. The CAI plasmids are isolated from these clones. The essential genes are identified by sequencing the genomic fragments in the isolated CAI plasmids.

This approach can also preferably be combined with a subtractive method (SCAI), one embodiment of which is depicted in fig. 4 for illustration. A genomic library with CAI plasmids (see fig. 1 and 3) is transferred into the homologous microorganism to be investigated, and the resulting individual clones are combined in a pool as bacterial CAI library. This pool is split up into two identical groups (the driver pool and the tester pool) for the selection.

The term "driver" is used in this connection for the pool of bacterial clones which is treated so that the inducible promoter is activated, and asRNA of the CAI vector is expressed. The "tester" pool contains an identical set of clones with CAI plasmids which, however, is kept under noninducing conditions and thus has wild-type properties.

Ordinarily, the "driver" pool is employed for the selection (for example in the animal), whereas the "tester" pool is stored untreated. However, it is also possible for both groups to be subjected to a selection, in which case only the "driver" pool is induced by addition of the signal (for example tetracycline). Clones in which the translation of an essential gene is inhibited through expression of a particular asRNA are lost from the group during the selection. After an appropriate time, the surviving clones in both groups are recovered and the CAI plasmids are isolated from the clones of both groups. The clones genomic fragments are then amplified by PCR, using oligonucleotide primers which hybridize with vector sequences which flank the cloned genomic fragments. The amplified DNA fragments which

represented parts of essential genes are enriched by subtractive hybridization (see fig. 8) and isolated.

A promoter suitable according to the invention for a CAI vector is, for example, the Tet promoter, whose activity can be controlled by a regulatory protein (in this case the Tet repressor) and can be induced by an extracellular signal (tetracycline). Further inducible promoters are known in the state of the art.

Antisense RNA-stabilizing elements are known to the worker skilled in this field and need not be explained in detail here.

The high efficiency of the CAI method in inactivating single genes in an organism derives from the overlapping cloning of small genomic fragments and from the synthesis, associated therewith, of different asRNA derivatives for a particular gene region. The probability of obtaining an asRNA which efficiently inhibits translation of the target gene which is sought is greatly increased in this way. Investigations of this type can be directed as a complete genome of a pathogen, which makes it necessary to examine a very large number of individual genomic fragments. Assistance from equipment (robots) is advantageous in this case in order to achieve a high sample throughput. However, it is possible in these cases to investigate only certain states, for example the growth of the cells in a particular medium.

It is possible by the additional use of subtractive method steps (subtractive conditional antisense inhibition, SCAI) to reduce, preferably greatly, the number of individual clones to be investigated.

Subtractive recombination mutagenesis (SRM) is preferably used for identifying facultatively essential genes. In contradistinction to the CAI method, permanent gene mutations are generated, and the enrichment of essential genes is achieved

by a subtractive step. The SRM method can, like the CAI method, be carried out with complete or partial gene libraries from pathogenic microorganisms.

The SRM method is based on the inactivation of individual genes in the genome of a pathogen by complete insertion of a particular suicide plasmid which is unable to replicate, or is able to replicate only under particular conditions such as, for example, a permissive temperature, in the organism to be investigated, this plasmid being part of a gene library. The plasmids are inserted into the genome by homologous recombination. Successful insertion is indicated by expression of a plasmid-encoded antibiotic resistance marker.

A preferred embodiment of the SRM method is illustrated in figures 5 to 8.

Figure 5 depicts a suitable SRM vector which, like the CAI vector, contains a genomic or subgenomic DNA fragment of the microorganism to be investigated, and an origin of replication (ori) for a cloning host (for example E. coli), one or more selectable marker genes and another conditionally active origin of replication for the microorganism to be investigated, for example a temperature-sensitive origin or an origin whose activity depends on a replication factor present *in trans* and which can be additionally introduced into the system. The fact that the SRM plasmid contains a genomic sequence of the microorganism to be investigated means that on transfection of this vector into this microorganism there is homologous recombination in which the complete SRM plasmid is inserted into the genomic gene of the microorganism, and the corresponding gene is, if it is one such, inactivated. This leads to an insertion mutant. Suitable inducible origins of replication are, as mentioned, temperature-sensitive oris or those which can be controlled by a factor such as, for example, the RGK factor pir or the pWV factor repA, which is added to the system *in trans*.

Insertion of an SRM plasmid (see fig. 5) into the genome of the microorganism to be investigated takes place by homologous recombination between the genomic

fragment, which has been cloned in the plasmid, of the microorganism and the complementary genomic DNA sequence. After the plasmid has been transferred into the appropriate microorganism, those clones in which the SRM plasmid is inserted into the genome are isolated via selection for the plasmid-encoded marker under nonpermissive conditions, that is to say with inactive replication. Excision of the SRM plasmid likewise takes place by homologous recombination. Replication of the inserted plasmid is initiated under permissive conditions, resulting in sufficient quantities of free plasmid in the cells so that the plasmid can be isolated again from the clone. If the SRM plasmid has been inserted into an essential gene, the viability of the relevant clone is restricted (A), whereas mutants in nonessential genes have normal viability (B).

Just as in the CAI method, it is possible to transfer a library of insertion plasmids from genomic fragments of the microorganism to be investigated into this microorganism and form genomic insertion mutants. This preferred embodiment of the SRM method is depicted in figure 7. A library of SRM plasmids which contain individual genomic or subgenomic fragments is transferred into the homologous microorganism to be investigated. Genomic insertion mutants are selected with the aid of a plasmid-encoded marker (see fig. 5) under conditions which do not permit plasmid replication, such as, for example, at a nonpermissive temperature. It is possible in this step for only insertion mutants which are mutated in a nonessential or facultatively essential gene to survive, since mutants of an essential gene are not viable. The individual insertion mutants are combined in a pool, and this pool is then divided into two identical groups, the driver pool and the tester pool. The driver pool is selected for example by infecting an animal. The tester pool remains untreated. The selection results in loss from the driver pool of those clones which contain an insertion in a facultatively essential gene (which is necessary for survival and reproduction under the selection conditions). The plasmids inserted into the genome of the microorganism are then recircularized under permissive conditions and recovered from the surviving clones of both pools. Plasmids which contain fragments of facultatively essential genes are

missing from the driver pool. The fragments cloned in the SRM plasmids are then amplified by PCR in both pools (see fig. 4). Those amplified DNA fragments which represent parts of facultatively essential genes are enriched by genetic subtraction (see fig. 8) and isolated.

A particular embodiment which makes use of subtractive hybridization for enrichment in fragments of essential genes is illustrated by way of example in figure 8.

A: PCR-based genetic subtraction. The tester DNA fragments (see fig. 4 and 7) are ligated to an adapter oligonucleotide in such a way that the adapter is covalently linked only to one of the two DNA strands of a double-stranded tester DNA fragment, which is achieved, for example, by ligating a double-stranded, nonphosphorylated adapter to the 3'-phosphorylated DNA fragment of the tester DNA. These tester DNA fragments are then mixed with a molar excess of driver DNA fragments. The mixture is denatured and slowly rehybridized. Protruding single-strand ends are then filled in to give the double strand with DNA polymerase. The products of this reaction are amplified by PCR, using oligonucleotide primers corresponding to the adapter sequences. Only those tested DNA fragments which have not hybridized with driver DNA fragments are exponentially amplified thus enriched and are subsequently isolated by cloning.

B: Genetic subtraction by physical removal of biotinylated DNA fragments. The driver DNA fragments are biotinylated and subsequently mixed in excess with tester DNA fragments, denatured and slowly rehybridized. The biotinylate is homo- driver-driver double strands and heteroduplexes (driver-tester double strands) are removed from the tester-tester homoduplexes by extraction with carrier-coupled streptavidin. The latter are isolated by cloning.

The insertion mutants, generated for example by SRM, are investigated for their changed biological properties in animal tests or cell culture systems. It is possible by deliberate use of specific host cells, for example cultivated macrophages, or host tissue, for example spleen, to select gene groups which determine essential properties of the pathogen, for example the colonization of particular host cells. If the surviving mutants are isolated from the cells, the mutants lack essential genes. Subtraction of the surviving mutants from the complete gene bank results in the mutants of the essential genes.

The CAI method or the SRM method is a very efficient method for unambiguous identification and characterization of essential genes. Since essential genes represent a natural target for inhibiting active ingredients, the described methods provide an ideal basis for developing novel active ingredients.

In the methods which are described hereinafter, the identified genes are employed directly for active ingredient screening, it being possible to dispense with elaborate purification steps by comparison with conventional methods. Bacterial carrier cells which can be employed for screening for prophylactic and therapeutic active ingredients are at the center of these methods.

The produced gene-deficient microorganisms are then tested for their growth ability or their survivability. Suitable test systems are, for example, *in vitro* systems, cell culture systems, tissue culture systems and animal models as natural environment. If the method is applied to *H. pylori*, the organisms are, on the one hand, grown on a so-called complete medium, the complete medium making the best possible preconditions for growth of *H. pylori* possible. At the same time, the deficient *H. pylori* organisms are grown in a culture which ought to correspond as accurately as possible to the natural environment of *H. pylori*. Used for this purpose are, on the one hand, cell cultures based on primary cultures or cell lines from gastrointestinal tissue or else differentiated primary tissue (spheroids) in culture medium. Further possibilities for simulating the natural environment of

H. pylori comprise the use of stimulated macrophages, since *H. pylori* has the ability not to be taken up and metabolized by the latter. It is additionally possible also to examine whether the deficient *H. pylori* organisms are able to establish themselves in immunodeficient mice over a certain period of time.

If a deficient *H. pylori* bacterium is in fact able to survive on complete medium but does not grow in a natural environment as described above, the gene which is deficient in this organism is referred to as a facultatively essential gene.

If the deficient *H. pylori* organism is able to survive in neither of the test habitats, then an obligately essential gene is involved.

It is generally possible to assign essential genes of microorganisms to one of these categories.

It is then possible from these results to identify the sequences suppressed in mutated or/by asRNA and in each case assigned into one of these two categories, or else to the category of nonessential genes if the gene-deficient organism shows no adverse effects on its growth ability.

A further object of the present invention is to provide a genetic method for isolating and cloning the identified essential genes from various clinical *Helicobacter* isolates or from heterologous pathogenic microbes of clinical importance. The method of the invention therefore further comprises the steps of

(v) producing primers for the amplification and detection of homologous gene sequences in heterologous microorganisms

(VI) identification of the homologous gene sequences.

A preferred procedure for these further steps in the method is to produce so-called megaprimers from the identified essential *Helicobacter* genes by PCR (polymerase chain reaction) whose sequence can be derived directly from the

corresponding plasmids of the mutagenized DNA sections. These primers can then be used to isolate the previously identified essential genes from various *Helicobacter* isolates. If these essential genes have correspondences in other microorganisms, it is also possible in some circumstances to use the primers for isolating these genes from microorganisms different from *Helicobacter.* A further possibility is then to determine the exact DNA sequence of the isolated genes and establish the genetic variance within the various *Helicobacter* isolates or between the various microorganisms. In the production of the megaprimers, DNA fragments with variable 3' ends result. Because of this property it is possible to employ the DNA fragments for isolating variable or related genes using the PCR method described.

Identification of specific active ingredients

For identifying novel immunological active ingredients from the pool of identified essential genes of a pathogen and for further development of these active ingredients, it is very effective to employ bacterial carriers because the identified genes can be cloned directly into these carrier systems and be expressed therein. Active ingredient screening then takes place directly with the aid of these recombinant bacterial carriers. The carriers preferably used are attenuated bacteria such as, for example, salmonellae, because they have a natural immunostimulant potential. If these attenuated bacteria are used as carriers or producers for the identified essential genes of the pathogenic microbes, and whether the mammal is immunized with these inoculation strains, it is possible to induce a persistent immune response.

The immunological properties of these bacterial carrier systems have now been refined so much that a targeted immune response can be induced (VanCott et al., 1998; carrier patent EP98116827. 1). This property is significant inasmuch as the various pathogens can often be effectively controlled only via a particular branch of the immune system. This means that protection-mediating antigens can be

identified only if they are presented in the correct form to the immune system. Only if the carrier used has been loaded with an effective antigen can a protective effect occur. The diverse immunological properties of bacterial carrier systems and their superiority to conventional synthetic adjuvants mean that they are particularly suitable for identifying immunologically relevant antigens.

In addition, the bacterial carrier systems can be equipped with efficient gene expression systems which allow even problematic antigens to be produced (PCT/EP91/02478, EP98116827.1). Because of the direct subcloning of the isolated essential genes and the simple manipulation of the bacterial carriers in the production and vaccination, it is possible for a large number of antigens to be tested for their immunogenic and protective potential in a relatively short time. In conventional methods by contrast, it is necessary for the test antigens to be subjected to time-consuming purification methods, and often difficulties arise even in the recombinant production of the selected antigens in bacteria, connected with the toxic effect of these antigens on the producing bacterial strain.

An important precondition for the development of active ingredients is to establish the immunogenic potential of the identified sequences in order to determine the extent to which the corresponding gene products are suitable for producing antibodies or vaccines.

To identify immunological active ingredients against clinically relevant *Helicobacter* organisms it is necessary first to establish the extent to which the gene product of the identified essential gene has immunogenic properties. This means that it must be established by experiment whether a humoral and cellular immune response directed against the original gene product of the pathogen can be induced with the antigen in a mammal. This rules out in any event those antigens which are not recognized by the immune system during a natural infection. On the contrary, it might be expected that, for example in chronically infected people, the immune response to protection-mediating antigens is suppressed or

is of such a nature that no protective effect eventuates. However, those antigens which are subject to great genetic variation and thus are scarcely accessible to an effective immune response must be precluded.

To demonstrate the identity of the identified gene product in a naturally occurring infection, antiserum is obtained from patients who are either suffering from an active gastritis with symptoms, or from patients in whom *Helicobacter* infection causes no symptoms. These sera are used to test the electrophoretically fractionated recombinant protein in a conventional Western blot method. If a recognition reaction with a recombinant protein takes place with each of the two sera, that is to say that from a patient with a fulminant and that from a patient with an asymptomatic helicobacter infection, then this protein is involved in a natural infection. If, on the other hand, the recombinant polypeptide is recognized only by the serum from the patient with an asymptomatic infection, this may additionally be evidence of a protective potential of the corresponding protein. It is moreover possible that antibodies specifically directed against this protein may be employed for passive immunization.

Also of special interest are antibodies from individuals demonstrated not to be *Helicobacter* carriers, because they suggest a protective potential for a corresponding recombinant polypeptide.

In addition, the immunogenic polypeptides are employed together with suitable additives for immunization *in vivo*. For this purpose, various adjuvants, bacterial toxins, cytokines or a polypeptide of the invention are used as live vaccine. The immune response is tested to find whether it elicits a protective effect against other homologous infections (for example infections with different *H. pylori* strains) after administration of a particular polypeptide of the invention in combination with appropriate additives after infection with the homologous microbe.

Yet a further possibility for testing the immunogenicity is to induce in an animal model (for example mouse or rabbit) an immune response to *Helicobacter* or other microorganisms, and to obtain from the immunized animals antibodies which can then be used in another Western blot analysis. At the same time, patients' biopsies must be immunologically investigated *in situ* with the same antibodies, because *Helicobacter* and other microorganisms in culture may lose or acquire particular proteins.

In parallel with this, preference must be given to investigation of whether a protective effect can be achieved against an infection with heterologous microbes (preferably other Gram-negative bacteria) which express the corresponding polypeptide.

After it has been found whether the identified genes or their expression products are able to elicit an immune response, further investigation by the method of the invention is possible into whether a pre-existent infection can also be treated with such antigens. If it is possible in this way to identify a polypeptide which shows a therapeutic effect, it will preferably also be investigated for its activity on infections with heterologous microbes.

Screening for immunological substances with prophylactic or therapeutic activity may take place in accordance with the following scheme, compliance with the individual steps not being obligatory:

1. Cloning of the identified gene into a suitable bacterial carrier strain and detection and quantification of the complete gene product by SDS-PAGE.
2. Immunological characterization of the generated gene product with the aid of (a) sera from infected or/and naturally protected hosts, which ought to recognize the gene product in the carrier strain; (b) hyperimmune sera from animals immunized with the recombinant carrier strain. Where the particular hyperimmune serum should recognize the original gene product in the

pathogenic microbe. It may be possible in this case for the original antigen to be produced by the pathogen only in a particular period of development.

3. The protective effect of the individual antigens in the prophylactic or/and therapeutic use form is investigated in an animal model.

All protective antigens identified by the described process can then be further developed in a second step. The priority in this further development is, inter alia, to evaluate the genetic constancy of the identified protective antigens within the pathogen and related pathogenic microbes in its distribution around the world. In addition, for the development of more effective vaccines, genetic differences in the immune system of the inoculated people are examined. The aim of both methods is to identify antigens or epitopes which have maximally broad applicability. The so-called megaprimer approach can be employed to detect genetic variability within a species or homologous microbes. Gene-specific primers with variable 3' ends are produced by PCR from the plasmid having the relevant gene and make it possible to amplify homologous genes. It is possible on the basis of the DNA sequence established for the amplified genes to deduce the variability thereof and, for example, determine genetically constant regions.

The genetic differences between inoculated individuals in the reaction to a defined antigen can be evaluated by means of *in vitro* vaccination. For this purpose, antigen-presenting cells (APC), for example dentritic precursor cells, are isolated from different donors and are expanded *in vitro* and mixed with the antigens to be tested, the antigens preferably being expressed by appropriate vectors. The identified genes can also be expressed singly or in defined combination in dendritic cells (DC) of uninfected donors. In this case, the gene products are processed by the host cell and presented by the MHC complex. DC are particularly suitable for antigen presentations to naive or "sleeping" T cells. Incubation of DC with T cells from autologous donors makes it possible to determine whether this donor would react to the antigen employed if he came into contact therewith in a natural way, for example through a protective vaccination. It is possible on the

basis of the immune response of the T cells to decide about a possible immunogenicity of the corresponding antigen. Such an immune response consists, for example, of a proliferation of T cells or of cytokine release, in particular IL-2 and IL-4. The cytokines can be analyzed, for example, with the aid of a commercially available assay kit (for example from Genzyme Cambridge M.A.).

Finally, the antigens or epitopes identified and characterized in the manner described can be employed for developing the first vaccine prototypes. A distinction is made in this connection between two types of vaccine, active and passive inoculation.

For passive immunization, antibodies or antibody fragments with a protective or inhibiting effect are supplied to the inoculated person from the outside.

Antibodies are provided in the form of polyclonal, but preferably monoclonal, antibodies (MAbs) or recombinant antibodies. These include antibodies which react specifically with polypeptides of the invention or their subunits and fragments and can be used for a prophylactic and/or therapeutic application, for example a passive immunization. These anti-protein or anti-peptide antisera or monoclonal antibodies can be produced by means of standard protocols, for example by immunization of animals such as mice, rats or goats with a purified polypeptide of the invention, a fusion protein or a subfragment thereof. In addition, the animals can also be immunized with bacterial vaccine carriers which are equipped with corresponding genes of the invention and express the encoded polypeptides. The antibodies in this case are preferably directed immunospecifically against antigenic determinants or epitopes thereto of the described *Helicobacter* polypeptides or a closely related polypeptide which has a homology of at least 90%. They do not cross-react with polypeptides which, for example, have a homology of less than 80%.

Starting from a cell line which produces a polypeptide-specific monoclonal antibody, it is possible to create from the gene encoding such an antibody chimeric genes which determine antibodies consisting of a mouse antigen-binding domain and the Fc part of a human antibody. These antibodies can be produced in cell lines or transgenic animals.

In place of antibodies generated in an animal, it is also possible to use antibody fragments, miniantibodies, which are produced, for example, in a heterologous system such as bacteria. These miniantibodies may be either monovalent or bivalent and consist of dimerized single-chain molecules (Kujau et al., 1998; Kalinke et al., 1996; Pack et al., 1993).

Antibodies against the immunogenic polypeptides of the invention can also be generated in plants. Examples thereof are described, for example, by Hiatt and Ma (1993), van Engelen et al. (1994) and Ma et al. (1994). According to the plant used in each case, these can be, for example, used directly for consumption and thus as oral vaccine.

Another very widely applicable way of producing antibodies is in the milk and eggs of immunized animals. If, for example, suitable antigens are administered to pregnant cows, sheep or horses, immunoglobulins which can be used to develop a vaccine are found in the milk. The milk can then either be administered directly as vaccine, or a concentrated immunoglobulin extract can be produced. It is also possible in the same way to produce antibodies (hyperimmune antibodies) in chicken eggs (Ling et al., 1998; Sasse et al., 1998). The described immunogenic polypeptides of the invention can therefore also be used for developing a dairy product or chicken eggs which can be used as oral vaccine.

The generated antibodies or fragments thereof are tested for their applicability according to the invention. For this purpose, they can be purified by known methods (precipitation, chromatographic methods) and, for example, investigated

to find whether they are able to inhibit the process *of H. pylori* infection (adhesion assays) or have activating effects on complement or ADCC ("antibody-dependent cell-mediated cytotoxicity").

For passive immunization, the antibodies generated with the aid of the polypeptides of the invention are administered either orally or intragastrically. For this purpose the antibodies are mixed with a bicarbonate buffer. However, they can also be administered systemically, in which case they need not be buffered.

Antibodies are preferably used alone or else in combination with other nonimmunological active ingredients, for example with antibiotics or proton blockers.

Active vaccination is based on an immune response induced by the inoculated organism itself. Preferred dosage forms of vaccines are as antigens, antigen fragments, subunit vaccine, as DNA vaccine, as live vaccine, or as food vaccine.

Antigens are those polypeptides or fragments thereof which are able to elicit an immune response *in vivo*.

If the polypeptide is to be provided as subunit vaccine, it is initially broken down into its subunits or structural domains in accordance with its antigenicity pattern (for example T- and B-cell epitopes). This antigenicity pattern can be constructed with the aid of a computer program, it being possible to recognize immunogenic regions consisting of a short polypeptide sequence of about 8 to 10 amino acids (Hughes *et al.*, 1992). The individual polypeptide pieces can then be tested for their immunogenicity in mice or in primates or humans. For this purpose, they can be administered either as purified polypeptides produced synthetically, in combination with appropriate additives such as an adjuvant, toxin or cytokine, or as fusion protein coupled to a protein or protein subunit which is known to be immunogenic, such as, for example, the cholera toxin B subunit (Liljeqvist *et al.*,

1997). The immunogenic peptides can moreover be incorporated into outer membrane proteins such as, for example, the OmpS maltoporin of *E. coli* and be expressed heterologously in a vaccine carrier strain (Lang and Korhonen, 1997).

To develop a DNA vaccine, the polynucleic acid molecules characterized in the invention can be administered "naked" fused to a eukaryotic tissue-specific promoter or in the form of a plasmid. The "naked" DNA or the corresponding plasmid is administered in combination with an additive such as a reagent which alters cellular permeability such as, for example, bupivacaine (WO94/16737), cationic lipids such as, for example, DOTMA (N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethyl-ammonium chloride, DOTAP (1,2-bis(oleyloxy)-3-trimethyl-ammonio)propane), DDAB (dimethyl-dioctadecyl-ammonium bromide), DOGS (dioctadecyl-amidolglycyl-spermidine) or cholesterinol derivatives, silica, gold or tungsten (Tang *et al.* 1992) or packaged in liposomes (WO93/18759, WO93/19768, WO94/25608, WO95/2397) or microparticles. Examples of promoters and gene ferries which can be used have been described by Hartikka *et al.* (1996). However, it is possible to use, for example, attenuated salmonellae for administering the polynucleotide molecules. The bacteria are for this purpose transformed with eukaryotic expression vectors which comprise a polynucleotide molecule of the invention, and then administered orally. The plasmid DNA is subsequently transferred from the bacterium to the host (Darji *et al.*, 1997). However, it is also possible to use filamentous phages for transforming attenuated carrier bacteria. The advantage of this is that they are able to transfer an extremely large number of plasmids.

Available for the development of a live vaccine are, inter alia, viral such as, for example, adenoviral or chickenpox virus vectors, and bacterial vectors such as, for example, *Salmonella, Shigella* or *Lactobacillus*. Attenuated, nonvirulent *Salmonella typhimurium* strains which can be used for the recombinant expression of heterologous antigens and be administered orally have been characterized many times (Mekalanos, 1994, WO92/11361, Cirillo *et al.*, 1995

and Dorner (1995). Further bacterial vectors which can be used as vaccine vectors has been described by Cirillo *et al.*, (1995) and Dorner (1995). A polynucleotide molecule of the invention which codes for a polypeptide with therapeutic or prophylactic activity is for this purpose either stably integrated into the bacterial genome and subjected to a transport system which makes presentation on the bacterial surface possible (PCT/EP94/04286; WO97/35022). The corresponding polynucleotide molecule may be present in the bacterium or else as plasmid in the free state.

Vaccines are usually administered with suitable additives such as, for example, adjuvants, bacterial toxins, cytokines etc. which assist the immunogenic polypeptide in its protective or therapeutic effect. Adjuvants with few side effects for use in humans which are suitable for subunit vaccines and live vaccines but also, in some cases, for DNA vaccines are, for example, aluminum hydroxide, aluminum phosphate, calcium phosphate, N-acetyl-muramyl-L-threonyl-D-isoglutamine, N-acetyl-normuramyl-L-alanyl-D-isoglutamine, N-acetyl-muramyl-L-alanyl-D-isoglutamyl-L-alanin-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxy-phosphoryloxy)ethylamine, liposomes, monophosphoryl-lipid A, trehalose dimicoloate, fungal polysaccharides such as, for example, schizophyllan, muramyl dipeptide, muramyl dipeptide derivatives, and phorbol esters, saponins and immunostimulating complexes (ISCOMS) (Gupta and Siber, 1995). It is possible to use as bacterial toxin for example cholera toxin or its subunits or the heat-labile toxin from *E. coli*. Although these are highly potent adjuvants, the possibility of using them in humans is limited because of their toxicity. However, molecules which are active but nontoxic can be developed with the aid of certain mutagenesis techniques (O'Hagan, 1998).

Another possibility to optimize the immune response to the substances of the invention with prophylactic and/or therapeutic activity is to express the corresponding polypeptide as fusion protein with an immunogenic protein domain. One possibility is, for example, to use the pilin DSL domain from *Pseudomonas*

aeruginosa as fusion partner. Fusion proteins fused to glutathione S-transferase or thioredoxin have also been described (Hill et al., 1997; Gabelsberger et al., 1997). The corresponding fusion protein can in each case be produced and administered as subunit or live vaccine.

The immune response of the identified immunologically active substances can also be modulated by administering them in combination with certain cytokines. Suitable for simultaneous administration is coexpression of the polynucleotide sequences of the invention with a particular cytokine in *Salmonella* or another host bacterium. The corresponding cytokine can for this purpose can be encoded either on a separate plasmid, in series or as fusion protein with the desired polynucleotide sequence of the invention and then be transformed into the host bacterium. Suitable cytokines are those like, for example, interleukin-6 (IL-6), interleukin-10 (IL-10) or interleukin-12 (IL-12) which stimulate the immune system.

For further optimization it is possible to combine individual substances of the invention with immunogenic activity with one another or in combination with known immunogenic substances such as, for example, VacA or its individual subunits. The corresponding polypeptide can thus be expressed together with at least one other *Helicobacter* antigen such as, for example, the native urease or its subunits, fragments, homologs, mutants or derivatives thereof. In addition, for example, various subunit vaccines can be administered singly or together as fusion protein as described hereinbefore. It is possible to use for this purpose in turn, for example, purified polypeptide molecules in combination with a suitable adjuvant, bacterial toxin or cytokine. A further possibility is for various combinations of nucleotide sequences of immunogenic subunits of the described polypeptide molecules to be produced on a common plasmid and administered as live vaccine. A vaccine vector of the invention may thus comprise one or more polypeptides of the invention, derivatives or fragments thereof. An additional possibility is to combine a DNA vaccine with one or more purified subunit vaccines in a suitable carrier substance as already described hereinbefore.

It is likewise possible to employ bacterial carriers very effectively for identifying novel pharmacological active ingredients from the pool of identified essential genes of a pathogen and for further development of these active ingredients, because the identified genes can be cloned directly into these carrier systems and expressed therein. Active ingredient screening then takes place directly with the aid of these recombinant bacterial carriers.

Display systems are used to present an expressed polypeptide of the invention on the bacterial cell surface. Transport of polypeptides through the inner membrane is made possible by a signal peptide on the amino end, while other portions undertake incorporation and anchoring in the outer membrane. Various outer membrane proteins of *E. coli* have been described as carrier portions, such as, for example, PhoE (Agterberg *et al.*, (1990) or OmpA (Francisco *et al.*, 1992). However, it is also possible to use fusions with the transport domain of the IgA protease precursor IgAß (Klauser *et al.*, 1990). Examples of display systems are, for example, the DsbA system (PCT/EP 94/02486) or the autotransporter system (AIDA; WO97/35022). The polypeptides presented on the surface can then be used for binding studies with peptide libraries or combinatorial chemical substance libraries. The binding studies can be carried out with the aid of a high throughput system which makes high test rates possible, in liquid or else bound form. For this purpose, the presented polypeptides are coupled, for example, to a chromatophore which, in combination with another chromatophore which is coupled to the active ingredient peptide or the chemical substance, makes a color reaction possible. The corresponding active ingredient component can, however, also for example be tagged with a fluorescent dye or coupled to a solid carrier matrix. On use of a solid phase system, either the polypeptide of the invention which is used or else, conversely, the peptide library or combinatorial active ingredient library is coupled to the carrier matrix beforehand. The particular dye-tagged component then binds to the immobilized component, again making a color reaction possible. After the binding reaction it is then necessary to carry out several washing steps before the corresponding substance is isolated. The

advantage of the solid phase system compared with binding in liquid is that the active substance can be isolated more quickly, because the unbound substances are washed away.

Further methods for identifying novel pharmacological active ingredients are based on knowledge of the primary structure of the identified essential genes or make use of the purified recombinantly produced gene products.

An alternative embodiment of the method of the invention consists of finding specific binding partners of the polypeptides encoded by the identified genes.

It is possible and preferred to carry out homology studies with *Helicobacter* and other organisms, for example by means of computer alignment, Southern blots, PCR and the like and subsequently assign the sequences. It is then possible, on the basis of homologies, to draw conclusions about potential binding partners of the proteins.

It is likewise possible and preferred to carry out combinatorial binding studies via target-directed screening methods with the aid of substance libraries. The potential binding substances are bound in a special arrangement, for example in microtiter plates or other carrier materials. The target, usually purified, where appropriate recombinant, *Helicobacter* proteins, is then added in a soluble form, which makes it possible to detect an interaction between the target and particular substances. Indirect detection is possible by labeled antibodies directed against the substance or by the introduction of additional elements (tags) into the target.

A further variant of binding studies consists of the expression of *Helicobacter* proteins in recombinant bacteria (for example those which produce the fluorescent protein GFP or particular enzymes) which present the proteins on the surface, and subsequent testing of a substance library.

A further possibility is to establish the three-dimensional structure of the polypeptides encoded by the genes identified according to the invention, or fragments thereof, by crystallographic analysis. If the resolution is sufficient, possible "pockets" or other binding sites in their three-dimensional structure can be accurately characterized. The structure of potential binding partners can be calculated on the basis of these data.

It is possible by use of methods such as, for example, a two-hybrid system, display systems, high throughput screening or combinatorial binding studies to identify randomly generated polypeptides which bind to the *Helicobacter* polypeptides of the invention or fragments thereof, or to other polypeptides identified according to the invention. If a corresponding peptide is found in this way, it can be further modified chemically until the optimal possible binding is reached. The identified polypeptide can be used, for example, as inhibitor by, for example, coupling it to a toxin and an internalization signal which destroys the pathogenic microbe, or else as peptide mimetic in order to prevent the binding of the microbe to the cellular surface (EP-412,762A and EP-B31,080A). The two-hybrid system can, however, also be used to generate activators of the immune system. The identified peptides which bind to the *Helicobacter* polypeptides of the invention can for this purpose be coupled to particular ligands, for example for the T-cell receptor. Thus, if these peptides equipped in this way are administered to an animal or human infected by pathogens, the body's own immune system is specifically attracted and activated.

Compliance with the individual steps is not obligatory in this connection; on the contrary, addition or substitution of other steps is possible.

The respective prototypes of an immunological or pharmacological active ingredient are subsequently developed further and improved.

Further development of a vaccine can take place by combining a plurality of antigenic gene products or parts thereof in an active ingredient and/or administering with various carriers or additives. Variously attenuated bacterial or viral organisms function as carriers, and adjuvants and/or cytokines function as additives.

For further development of a pharmacological active ingredient, a lead structure which has been characterized as effective is further modified chemically so that optimal binding and inhibition of the identified gene product takes place. In addition, the active ingredients ought to be well tolerated by the patient and have few side effects.

For identifying active ingredients which bind to polynucleotides it is possible for a polynucleotide of the invention to be previously coupled for example to a carrier matrix, or vice versa, the active ingredients of the polypeptide library or combinatorial substance library. The subsequent scheme which is used is the same as that already described above for the polypeptides.

Such inhibitory substances may be polypeptides, peptides, but also chemical substances such as, for example, antibiotics. The inhibitory effect may moreover intervene in various stages of the replication of the microorganisms to be controlled. Examples are expression inhibitors or enzyme inhibitors or other inhibitors which are able to influence the natural function of the polypeptides of *Helicobacter* and related microorganisms. The invention likewise relates to such inhibitory substances.

Another possible way of finding an optimal active ingredient for *Helicobacter* and other bacterial infections is with the aid of special computer programs. It is possible on the basis of crystallographic data obtained for polypeptides described in the invention to produce a model which links together steric, electronic, hydrophobic and so-called resulting binding moments (RBMs) (Ray et al., 1998).

It is possible by using this model subsequently to model substances on the computer which, although they may contain the previously identified lead structures, are equipped with better binding properties. However, completely novel active ingredients may also be designed.

A further aspect of the present invention is a nucleic acid which codes for an essential secretory gene from *Helicobacter pylori* which has been identified by the method of the invention described above. Essential *Helicobacter* genes have been identified using the present method, and their nucleic acid sequences are indicated in SEQ ID NO. 1 to 245 (odd numbers). A nucleic acid of the invention is, for example, characterized in that it comprises

(a) one of the nucleic acid sequences depicted in SEQ ID NO: n, where n is an odd integer from 1 to 245 inclusive, or a protein-encoding section thereof,

(b) a nucleotide sequence corresponding to one of the sequences from (a) within the scope of the degeneracy of the genetic code or

(c) a nucleotide sequence hybridizing with one of the sequences from (a) and/or (b) under stringent conditions.

Besides the nucleotide sequences of the invention shown in the sequence listings, and nucleotide sequences corresponding to these sequences within the scope of the degeneracy of the genetic code, the present invention also comprises nucleotide sequences which hybridize with one of the aforementioned sequences. The term "hybridization" according to the present invention is used in Sambrook et al. (Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989), 1.101-1.104). Reference is preferably made to stringent hybridization when a positive hybridization signal is still observed after washing for one hour with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, particularly preferably at 62°C and most preferably at 68°C, in particular for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, particularly preferably at 62°C and most preferably at 68°C. A nucleotide sequence hybridizing under such washing conditions with one or more of the nucleotide sequences of the invention or with

a nucleotide sequence corresponding to these sequences within the scope of the degeneracy of the genetic code is a nucleotide sequence of the invention.

The nucleotide sequence of the invention is preferably a DNA. However, it may also comprise an RNA or a nucleic acid analog such as, for example, a peptidic nucleic acid. The nucleic acid of the invention particularly preferably comprises a protein-encoding section of the nucleotide sequences depicted in the sequence listing or a sequence which has a homology of more than 80%, preferably more than 90% and particularly preferably more than 95% with the depicted nucleotide sequences or a section thereof which is preferably at least 20 nucleotides (nt) and particularly preferably at least 50 nt long.

The homology is indicated in percent of identical positions on comparison of two nucleic acids (or peptide chain), where 100% homology means complete identity of the compared chain molecules (Herder: Lexikon der Biochemie und Molekularbiologie, Spektrum Akademischer Verlag 1995).

A nucleic acid of the invention may code for a secreted polypeptide with signal peptide or for a secreted polypeptide without signal peptide.

A nucleic acid of the invention comprises both the sequence of the coding strand and the sequence complementary thereto. The latter may be used, for example, for producing antisense nucleic acids.

The invention likewise of course also relates to a gene library comprising at least 2, preferably at least 20, more preferably at least 100 of said nucleic acids cloned in vectors.

Tables I and II indicate a list of the nucleic acids of the invention indicated in the sequence listing, together with their gene products and the functions and putative functions thereof.

The genes indicated in the sequence listing in their nucleic acid and amino acid sequences are bacterial genes. Apart from the AUG codon (mRNA), prokaryotes also use other (alternative) start codons. These are AUU (normally codes for isoleucine, Ile), UUG (normally codes for leucine, Leu) and GUG (normally codes for valine, Val). In the sequence listing, the amino acid sequences have been translated in accordance with the translation code normally used. It is pointed out that when reading the sequence listing the prokaryotic use of alternative start codons must be taken into account, and thus that the amino acid sequences which are encoded by the nucleic acid sequences shown in SEQ ID NO. 35, 49, 61, 69, 75, 81, 103 and 105 and which start with a methionine residue (Met) in place of Leu, Val or Ile residues respectively depicted in the sequence listing are also disclosed herein.

The following gene sequences also start with alternative start codons: **HPC152**(UUG); **HPC190**(UUG); **HPN048**(GUG); **HPN132**(GUG); **HPC010**(GUG); **HPC036**(UUG); **HPC056**(UUG); **HPC161**(GUG). These codons are usually translated as Met when they function as start codons.

Table I: Obligately essential genes

SEQ ID NO	Gene (obligately essential)	Length	Possible properties of the gene products	Signal peptide
1+2	HPS001 (HPS166)	873 bp	Signal peptidase I	+
3+4	HPC003	267 bp	Protein of flagella biosynthesis	-
5+6	HPC005	714 bp	Lipoprotein	+
7+8	HPC029 (HPC030)	552 bp	Secreted protein	-
9+10	HPS042	858 bp	Inner membrane protein, ubiquinol oxidoreductase	-
11+12	HPC057 (HPC109, HPC138)	192 bp	Secreted protein	+
15+16	HPS065 (HPS153)	1629 bp	Lipoprotein	+
17+18	HPS066	1377 bp	Inner membrane protein, iron-sulfur binding protein	-
21+22	HPS074	957 bp	Secreted protein	-
23+24	HPS083	480 bp	Secreted/periplasmic protein	-
25+26	HPS084	1983 bp	Secreted/periplasmic protein	+
27+28	HPC085	372 bp	Secreted/periplasmic protein	+
29+30	HPC090	558 bp	Secreted/periplasmic protein	+
31+32	HPS104	768 bp	Integral membrane protein	-
33+34	HPS115	2367 bp	ATPase for cation transport	-
35+36	HPS120	2751 bp	Outer membrane protein, protective surface antigen	+
37+38	HPS130	990 bp	Secreted protein	+
39+40	HPS133	1482 bp	Murein precursor protein	-
41+42	HPC134	600 bp	Inner membrane protein, protein translocation protein	-
43+44	HPS143	1536 bp	Secreted/periplasmic protein	+
45+46	HPS144	540 bp	Peptidoglycan-associated lipoprotein	+
49+50	HPS152	1062 bp	Secreted/periplasmic protein	+
51+52	HPS155	2202 bp	Protein of flagella biosynthesis	-

53+54	HPC157 (HPC181)	189 bp	Secreted protein	-
55+56	HPS183	1008 bp	Iron(III) ABC transporter	+
57+58	HPS186	240 bp	Secreted protein	-
59+60	HPS188	1764 bp	Outer membrane protein	+
61+62	HPS190	1443 bp	Secreted/periplasmic protein	-
115+116	HPN165	8709 bp	Toxin-like outer membrane protein; autotransporter	+
117+118	HPC001	873 bp	Signal peptidase I	+
119+120	HPC042	663 bp	Inner membrane protein, ubiquinol oxidoreductase	-
121+122	HPC065	1674 bp	Lipoprotein	+
123+124	HPC066	697 bp	Inner membrane protein, iron-sulfur binding protein	n.d.
125+126	HPC074	519 bp	Secreted protein	n.d.
127+128	HPC083	480 bp	Secreted/periplasmic protein	-
129+130	HPC084	1983 bp	Secreted/periplasmic protein	+
131+132	HPC104	768 bp	Integral membrane protein	-
133+134	HPC115	1510 bp	ATPase for cation transport	-
135+136	HPC120	1017 bp	Outer membrane protein, protective surface antigen	+
137+138	HPC130	193 bp	Secreted protein	+
139+140	HPC133	530 bp	Murein precursor protein	-
141+142	HPC143	1536 bp	Secreted/periplasmic protein	+
143+144	HPC144	88 bp	Peptidoglycan-associated lipoprotein	+
145+146	HPC152	1080 bp	Secreted/periplasmic protein	+
147+148	HPC155	695 bp	Protein of flagella biosynthesis	n.d.
149+150	HPC165	384 bp	Toxin-like outer membrane protein autotransporter	+
151+152	HPC183	1008 bp	Iron(III) ABC transporter	+
153+154	HPC186	240 bp	Secreted protein	-
155+156	HPC188	264 bp	Outer membrane protein	+
157+158	HPC190	1443 bp	Secreted/periplasmic protein	-

Table II: Facultatively essential genes

SEQ ID NO	Gene (facultatively essential)	Length	Possible properties of the gene products	Signal peptide
99+100	HPS004 (HPS027, HPS121, HPS131)	1644 bp	Inner membrane protein	+
63+64	HP008 (HPC114, HPC145)	543 bp	Secreted/periplasmic protein	-
65+66	HPS013	1746 bp	2',3'-cyclic nucleotide 2'-phosphodiesterase	+
67+68	HPS024 (HPS025)	1698 bp	Chemotaxis protein	+
69+70	HPS036	855 bp	Secreted protein	+
71+72	HPS038	669 bp	Secreted protein	-
73+74	HPS039 (HPS147)	804 bp	Secreted protein	-
75+76	HPS040	1545 bp	Secreted protein	+
77+78	HPS048	912 bp	Secreted protein	+
79+80	HPS050	834 bp	Periplasmic glutamine-binding protein	+
81+82	HPS052	1296 bp	Outer membrane protein	-
83+84	HPS056	1197 bp	Secreted protein	-
85+86	HPS059	1131 bp	Integral membrane protein	+
87+88	HPS063	516 bp	ATP synthase F0, subunit b	+
89+90	HPS069	990 bp	Secreted protein	+
91+92	HPS091	684 bp	Integral membrane protein	-
93+94	HPS095	729 bp	Outer membrane protein	+
95+96	HPS099	975 bp	Secreted protein	-
97+98	HPS117 (HPS118, HPS162)	1290 bp	Secreted protein	+
101+102	HPS132	3063 bp	Cation efflux protein	+
47+48	HPC140 (HPC150, HPC179)	1557 bp	AlpA adhesin	+

103+104	HPS149	2028 bp	Methyl-accepting chemotaxis protein	-
105+106	HPS161	273 bp	Secreted protein	+
107+108	HPS176	759 bp	Protein of the cag pathogenicity island	-
109+110	HPS187	1245 bp	Zinc-dependent metalloprotease	+
111+112	HPS189	1566 bp	Secreted protein	-
113+114	HPS191	1782 bp	Secreted protein	+
13+14	HPS062 (HPS171)	957 bp	Secreted protein	-
19+20	HPS068	1533 bp	Lipase	+
159+160	HPN013	1401 bp	Secreted/periplasmic protein	+
161+162	HPN048	2577 bp	Cell division protein	+
163+164	HPN091	1329 bp	Sodium- and chloride-dependent transporter	-
165+166	HPN132	1785 bp	Periplasmic oligopeptide ABC transporter	+
167+168	HPN137	2007 bp	Secreted protein	+
169+170	HPN172	771 bp	Secreted protein	+
171+172	HPC004	1641 bp	Inner membrane protein	+
173+174	HPC010	783 bp	Secreted protein	+
175+176	HPC012	1131 bp	Secreted protein	+
177+178	HPC013	675 bp	Secreted/periplasmic protein	+
179+180	HPC024	348 bp	Chemotaxis protein	+
181+182	HPC034	1359 bp	2',3'-cyclic nucleotide 2'-phosphodiesterase	+
183+184	HPC036	858 bp	Secreted protein	+
185+186	HPC039	804 bp	Secreted protein	-
187+188	HPC048	1657 bp	Cell division protein	+
189+190	HPC050	684 bp	Secreted protein involved in flagellar motility	+
191+192	HPC056	879 bp	Hypothetical protein	+
193+194	HPC059	1131 bp	Integral membrane protein	+
195+196	HPC063	1068 bp	Secreted protein	-
197+198	HPC068	1533 bp	Lipase	+
199+200	HPC069	516 bp	ATP synthase F0, subunit b	+

201+202	HPC070	565 bp	Cation efflux protein	+
203+204	HPC076	663 bp	Secreted protein	-
205+206	HPC091	454 bp	Sodium- and chloride-dependent transporter	-
207+208	HPC094	921 bp	Secreted protein	+
209+210	HPC095	1572 bp	Conserved hypothetical integral membrane protein	+
211+212	HPC099	966 bp	Secreted protein	-
213+214	HPC101	812 bp	Periplasmic glutamine-binding protein	+
215+216	HPC107	1268 bp	Outer membrane protein	-
217+218	HPC110	312 bp	Secreted protein	n.d.
219+220	HPC117	1290 bp	Secreted protein	+
221+222	HPC129	471 bp	Secreted protein	+
223+224	HPC132	186 bp	Secreted protein	+
225+226	HPC137	1096 bp	Secreted protein	+
227+228	HPC149	1722 bp	Methyl-accepting chemotaxis protein	-
229+230	HPC161	273 bp	Secreted protein	+
231+232	HPC169	183bp	Integral membrane protein	-
233+234	HPC172	543 bp	Secreted protein	n.d.
235+236	HPC174	729 bp	Outer membrane protein	+
237+238	HPC176	540 bp	Protein of the cag pathogenicity island	-
239+240	HPC180	864 bp	Secreted protein	+
241+242	HPC187	1072 bp	Zinc-dependent metalloprotease	n.d.
243+244	HPC189	357 bp	Secreted protein	+
245+246	HPC191	1251 bp	Secreted protein	n.d.

n.d. = not determined

SEQ ID NO: 247: shows the sequence of the vector pSRM4 (figure 13).

The sequences SEQ ID No. 1 to 114 are depicted in the appended sequence listings.

The sequences SEQ ID No. 115 to 246 are depicted in figures 14 and 15.

Tables A and B below list the nucleic acid sequences of the identified helicobacter genes (gene ID) of which the nucleic acid sequences could not be established completely. The depicted reading frame of the established helicobacter gene sequences was verified on the basis of the selection after functional fusions with the indicator gene used, of β-lactamase (see example 3 and 4).

Table A lists the helicobacter genes which contain an incomplete 5' end and/or 3' end. Table B lists those helicobacter genes to which one or more amino acids cannot be assigned because of missing sequence data. Stop codons can be ruled out in all cases.

Table A Gene sequences with incomplete 5' and/or 3' end

Category: Obligately essential genes

Gene ID (internal)	Missing 5' end	Missing 3' end
HPC042		X
HPC066	X	X
HPC074	X	X
HPC115		X
HPC120		X
HPC130		X
HPC133		X
HPC144		X
HPC155	X	X
HPC165		X
HPC188		X

Category: Facultatively essential genes

Gene ID (internal)	Missing 5' end	Missing 3' end
HPC010		X
HPC024		X
HPC034		X
HPC048		X
HPC070		X
HPC076		X
HPC091		X
HPC101		X
HPC107		X
HPC110	X	X
HPC129		X
HPC137		X
HPC149		X
HPC169		X
HPC172	X	
HPC176		X
HPC187	X	X
HPC191	X	X

Table B Incomplete amino acid sequences

Category: Obligately essential genes

Gene ID (internal)	AA position[1]	NA position[2]
HPC001	26	73
HPC065	360	1078
HPC066	112	334
HPC074	164; 185	490; 493
HPC084	403; 500	1207; 1498
HPC104	204	610
HPC155	14	41

Category: Facultatively essential genes

Gene ID (internal)	AA position[1]	NA position[2]
HPC012	190; 235	568; 703
HPC013	203; 207; 219	607; 619; 655
HPC036	31	91
HPC048	68; 397	202; 1189
HPC056	6; 7	16; 19
HPC068	381; 386	1141; 1158
HPC069	144; 148; 155; 168; 169	430; 442; 463; 502; 506
HPC076	176;	526
HPC101	133; 162	397; 484
HPC107	221; 291	661; 871
HPC137	327	979
HPC149	279	835
HPC180	181; 191; 209	541; 571; 825

[1] Position of the missing amino acid in the derived amino acid sequence

[2] Position in the nucleic acid sequence, first base from the coding base triplet.

A further aspect of the present invention is a vector which comprises a nucleic acid of the invention or a section thereof. The nucleic acid or the nucleic acid section may be cloned into the vector in such a way that it can be expressed either in the sense or antisense direction. The nucleic acid section preferably has a minimum length of 15 nucleotides, more preferably 20 nucleotides, more preferably 50 nucleotides. This vector may be any prokaryotic or eukaryotic vector on which the DNA sequence of the invention is present preferably in conjunction with expression signals such as, for example, promoter, operator, enhancer etc. Examples of prokaryotic vectors are chromosomal vectors such as bacteriophages (for example, bacteriophage λ) and extrachromosomal vectors such as, for example, plasmids, with particular preference for circular vectors. Suitable prokaryotic vectors are described, for example, in Sambrook *et al.*, Molecular Cloning (1987), chapter 1-4. On the other hand, the vector of the invention may also be a eukaryotic vector, for example a yeast vector or a vector suitable for higher cells (for example a plasmid vector, viral vector, plant vector). Vectors of these types are described, for example, in Sambrook et al., supra, chapter 16. The invention likewise relates to CAI and SRM vectors as described above.

The present invention further relates to a cell transformed with a vector of the invention or with a nucleic acid of the invention. In a preferred embodiment, this cell is a prokaryotic cell, preferably a Gram-negative bacterium, for example *E.coli*. On the other hand, however, the cell of the invention may also be a eukaryotic cell such as, for example, a fungal cell, a yeast cell, an animal or a plant cell. The cell is particularly preferably a microorganism, for example *Helicobacter* or salmonellae. Microorganisms transformed with CAI or SRM vectors have already been described above. The invention likewise relates to them as well as to mutant libraries which can be produced using the above methods.

A further aspect of the invention relates to an essential and preferably secreted polypeptide of *H. pylori*. This is, in particular, a polypeptide which comprises

(a) one of the amino acid sequences depicted in SEQ ID NO: m, where m is an even integer from 2 to 246 inclusive, or

(b) a sequence which cross-reacts immunologically with one of the sequences according to (a).

Sequences which cross-react immunologically thus also include muteins, variants and fragments of the sequences depicted in SEQ ID NO. 2 to 246. By these are meant sequences which differ from the above sequences by substitution, deletion and/or insertion of individual amino acids or short amino acid sections.

It was possible on the basis of homology analyses using the FASTA protein program for particular features or a putative location in the bacterium to be assigned to the identified polypeptides whose sequence could be found using the nucleic acid sequence. Some of the polypeptides encoded by the nucleic acids of the invention have a signal peptide and are exported by the Sec-dependent transport mechanism to their target site, whereas others have no signal peptide and are therefore probably secreted via a Sec-independent transport mechanism, for example by the ABC transporter system (see tables I and II).

The invention further relates to a method for producing the polypeptides and fragments thereof of the invention. Polypeptides of the invention are preferably produced by transforming a cell with a DNA molecule or vector of the invention, cultivating the transformed cell under conditions with which expression of the polypeptide takes place, and isolating the polypeptide from the cell or/and from the culture supernatant. It is moreover possible for the polypeptide of the invention to be obtained both as fusion polypeptide and as nonfusion polypeptide.

The polypeptide of the invention can be used as immuncgen for producing antibodies.

The present invention thus also relates to an antibody directed against a polypeptide of the invention. The invention likewise relates to fragments of such antibodies such as, for example, Fab fragments or Fc fragments.

Yet a further aspect of the invention is an inhibitor of the polypeptides of the invention, fragments thereof, or expression, presentation or/and natural function thereof. This is preferably a molecule able to bind specifically to a polypeptide or fragment thereof or/and influence the expression, presentation or/and natural function thereof. The identification of such specific binding partners has already been described above. Particularly suitable inhibitors are proteins or peptides which inhibit the natural function of a polypeptide of the invention, for example enzymes can be inhibited by blocking the active site.

Yet a further aspect of the present invention relates to a pharmaceutical composition which comprises as active ingredient a DNA molecule of the invention, a vector of the invention, a cell of the invention, a polypeptide of the invention, an antibody of the invention or fragment thereof or/and an inhibitory molecule able to bind specifically to a polypeptide of the invention, where appropriate together with conventional pharmaceutical excipients, diluents, additives and carriers.

A pharmaceutical composition of the invention can be used in various ways and with use of individual components thereof as active substances for inhibiting the reproduction of *Helicobacter* organisms in a host, specifically in humans.

The pharmaceutical composition of the invention can be used on the one hand for diagnosing a *Helicobacter* infection. Diagnosis takes place at the nucleic acid level preferably by using hybridization probes or primers which have a specific DNA sequence which is complementary to at least one section of one of the sequences depicted in SEQ ID NO. 1 to 245 (odd numbers), so that they allow amplification of the sequences of the invention. As already mentioned, the

amplification primers or probes can also be used for the amplification and thus for the detection of related microorganisms if the latter have gene sequences which code for the same essential gene. Diagnosis at a protein level preferably takes place with the aid of the antibodies of the invention.

The pharmaceutical composition is moreover suitable for the prophylaxis and control of *Helicobacter* infections and infections with related microorganisms.

Another important aspect of the present invention is the use of the identified essential genes of *Helicobacter pylori* for the prevention or control of an infection with *Helicobacter* or related microorganisms. These identified essential genes can be used in particular for producing vaccines (see above).

Another use of the polypeptides of the invention is in the purification of antibodies against *H. pylori* polypeptides and against corresponding polypeptides from related and other microorganisms..

The invention is explained in more detail by the following examples, figures and the sequence listing. The nucleic acid and amino acid sequences listed in tables I and II are depicted in the sequence listing.

Figure 1 shows a diagrammatic representation of a CAI vector.

Figure 2 shows a diagrammatic representation of a method of conditional antisense inhibition(CAI).

Figure 3 shows a diagrammatic representation of the investigation of the viability of deficient microorganisms on the basis of their survival rates using the CIA method.

Figure 4 shows a diagrammatic representation of the subtractive CAI method (CAI).

Figure 5 shows a diagrammatic representation of an SRM vector.

Figure 6 shows a diagrammatic representation of the reversible inactivation of a gene by the insertion/excision of conditionally replicating SRM plasmid.

Figure 7 shows a diagrammatic representation of an SRM method.

Figure 8 shows a diagrammatic representation of the enrichment of fragments of essential genes by subtractive hybridization.

Figure 9 shows a survey map of pSRM4. The origin of replication (ori) is depicted as a bar, and coding regions (see below for detailed explanations) are depicted as arrows. Unique restriction cleavage sites are indicated.

Figure 10 shows a diagrammatic representation of the integration and excision of a pSRM plasmid which harbors an insert with homology to a gene of the recipient organism. The double arrows show the equilibrium reactions between integration and excision. Broken lines indicate the DNA region in which homologous recombination takes place.

Figure 11 shows a diagrammatic survey of the steps for constructing a random fragment library.

Figure 12 shows the complete nucleotide sequence and amino acid sequence of the gene fusion of the HPC001 protein with the RGS His^4 epitope. The vector sequences are shown white on black.

Figure 13 shows the nucleotide sequence of the SRM vector pRSM4.

Figures 14 and 15 show the nucleotide and amino acid sequences of other essential helicobacter genes.

Examples

Example 1

Identification of essential genes by producing gene-deficient microorganisms by subtractive recombination mutagenesis (SRM)

The application of the described overall method is described by way of example for the example of the pathogen *Salmonella*, but can also be carried out with other pathogenic microbes, for example helicobacter.

The SRM method is carried out with the vector pSRM4 which was specifically developed for this method (see figure 9, for sequence see SEQ ID No.: 247

The plasmid has the following properties:

pSRM4 harbors a temperature-sensitive replication function (*repA*$_{ts}$, ori+) from the *Streptococcus* plasmid pWV01 (Kok, *et al.*, 1984; Maguin, *et al.*, 1992) and replicates at the permissive temperature (30°C) in Gram-negative bacteria (*E. coli*, *Salmonella*, etc.) and in Gram-positive bacteria (*Bacillus*, *Lactococcus*, *Streptococcus*,etc.). Because of the mechanism of replication (rolling circle

mechanism), it is possible that the plasmid replicates just as well in Helicobacter pylori as in mycobacteria (del Solar, *et al.*, 1993; Kleanthous, *et al.*, 1991).

To test the stability of pSRM4 at 30°C in *E. coli* or *Salmonella*, the bacteria are grown in liquid medium (liqmed) with tetracycline (Tet) (17.5 µg/ml) at 30°C overnight, diluted 1:100 in [lacuna] with Tet, after incubation at 30°C for 3 hours again diluted $1:10^3$ in liqmed (t0) and then incubated without Tet at 30°C. At various times, aliquots of the culture are taken and plated out in suitable dilutions on solid medium (solmed) with and without Tet. After incubation at 30°C, the colony counts on both media are determined and the number of clones per ml of culture is calculated. The number of Tet-resistant and Tet-sensitive clones is identical for up to 7 hours after t0, and after 19 hours about 10% of the clones are still resistant to Tet.

pSRM4 replicates neither in *Salmonella* nor in *E. coli.* at 37°C.

The experiment is carried out as under (1) with the difference that the incubation takes place at 37°C (without Tet). After 80 to 180 minutes after t0, the number of Tet-sensitive clones starts to increase logarithmically, while the number of Tet-resistant (plasmid-containing) clones remains the same.

pSRM4 codes for a tetracycline resistance which confers resistance to 17.5 µg/ml tetracycline. The tetracycline resistance gene can be replaced by other resistance genes. Several restriction cleavage sites are available inside the vector for this purpose (see fig. 9).

pSRM4 contains a so-called multiple cloning site (mcs) with different unique recognition sequences for restriction enzymes.

The mcs is located in part of the *lacZ* gene which can be used for alpha complementation in *E. coli* strains such as DH5α, XL1-Blue, or EC101 (Law, *et al.*, 1995). It is possible in this way to identify plasmids harboring cloned DNA fragments in the mcs through blue/white differentiation on solid medium with X-Gal

(80 μg/ml) and IPTG (420 μM IPTG). Clones of EC101 harboring plasmids with inserts form white colonies, whereas colonies of clones containing plasmids without insert are blue.

pSRM4 can be transformed with an efficiency of $2x10^7$ to $2x10^8$ per μg of DNA in *E. coli* DH5α and EC101, and with an efficiency of $5x10^6/\mu$g in *S. typhimurium* ATCC14028. The transformation takes place by electroporation in accordance with standard protocols.

pSRM4 does not recombine with the genome of *S. typhimurium*.

pSRM4 was transformed into *S. typhimurium* ATCC14028; the transformants were cultivated on solmed with Tet at 30°C for 48 h and isolated. A single colony (about 10^8 cells) was resuspended in one ml of liqmed and incubated in appropriate dilutions on solmed with Tet at 30°C to determine the total cell count and at 37°C to determine the number of integrants. Among $1x10^8$ cells of the initial colony, one clone which forms colonies at 37°C was found. The frequency of nonhomolgous recombination of the plasmid with the genome of *S. typhimurium* is accordingly less than $1x10^{-8}$.

pSRM4 with short genomic inserts is suitable for the inactivation of genomic genes via homologous recombination (examples: *phoP* gene and *phoN* gene from *S. typhimurium*).

Internal fragments of the genes *phoP* (561 bp) and *phoN* (543 bp) were cloned into pSRM4. The corresponding plasmids (pSRM4-*phoP* and pSRM4-*phoN* respectively) were transformed into *S. typhimurium* ATCC14028 and the resulting clones were selected with Tet at 30°C. Two colonies were plated out in several dilutions on BCiP-containing minimal medium plates and incubated at 30°C and at 37°C. Colonies which grow at the nonpermissive temperature (37°C) consist of cells with genomically integrated vectors. These colonies show no or only a weak blue coloration because, owing to the insertion of the vector pSRM-*phoN* or

pSRM4-*phoP*, the gene *phoN* which codes for acid phosphatase PhoN and its activator gene *phoP* are inactivated and thus no reaction of the dye BCiP can take place. The average recombination rates calculated from the results are about $3.6x10^{-4}$. These recombination rates had also been observed for shorter genomic fragments with a length of up to 100 bp.

Clones growing at 37°C are examined for homologous recombination and integration of the plasmid into the genome. To detect integration, PCR reactions are carried out with a PCR primer which hybridizes with sequences of the plasmid, and a *phoN*- or *phoP*-specific primer which is homologous to a chromosomal *phoN* or *phoP* sequence but not to the cloned *phoN* or *phoP* fragment (primer pairs phoN for [GCTGTCGACTTTCTACCACTGATCGTAGC] / lacZ2 [CATGCCATGGCTGCGCGTAACCACC] and phoP3 [CCCCAAAGCACCATAATCAACGC] / lacZ1 [CATGCCATGGAAGAGCGCCCAATAC]). DNA of the plasmid is amplified together with a piece of chromosomal DNA which is located adjacent to the insertion site. On genomic integration of the plasmid, large amounts of PCR fragment are obtained. A PCR with the primers lacZ1 / SRMgb2 [ATACCGTCGACCTCGAG] is used to examine the loss of free plasmid DNA. Both primers hybridize only with sequences of the plasmid which flank the cloned insert; with genomically integrated plasmid, this amplification affords only a very small amount of product.

Genomically integrated pSRM4 excises at permissive temperature. Cloned genomic inserts can easily be prepared by PCR with various primer pairs(SRMgb2 / SRMgb4 [AACAAAAGCTGGGTACC]; lacZ1 / SRMgb2).

Integrant strains(*S. typhimurium* ATCC14028 derivatives with integration of pSRM-*phoN*) are incubated on solmed with Tet at 30°C and at 37°C and then plated out colony-wise on solmed with or without Tet, using various dilutions at both temperatures, in order to quantify the excision frequency. The quotient of the number of colonies on solmed without Tet and that on solmed with Tet indicates

the number of cells in a colony in which the plasmid is in free form compared with cells with genomically integrated plasmid. The quotient is a measure of the excision rate. Excision rates of from 3.2 to 17.1 result, depending on the initial colony.

Excized plasmid is in turn detected using the primers SRMgb2 and SRMgb4 or the primers lacZ1 and SRMgb2 in a PCR. Large amounts of PCR product result.

The insertion and excision of pSRM4 *fragment* via homologous recombination is an equilibrium reaction (see fig.10). At 30°C, the equilibrium is shifted toward plasmid freely replicating in the cell, and at 37°C mainly genomically integrated plasmid is present. Free and integrated plasmid can be detected by PCR both at 30°C and at 37°C. However, the respective ratios of amounts are shifted according to the temperature sensitivity of the plasmid.

A mutant containing pSRM4-phoP in the *phoP* gene is stable and attenuated in a cell culture system.

To examine the stability and attenuation of a pSRM4-*phoP* mutant, the intracellular reproduction of this mutant in murine J774A.1 macrophages is compared with the reproduction of pSRM4-*phoN*, with the wild-type *S. typhimurium* ATCC14028, and with stabile *phoP* and *phoN* mutants. For this purpose, $4x10^5$ macrophages of the cell line J774A.1 are infected with $8x10^6$ bacteria of a clone at 37°C, 5% CO_2. After 30 min, the extracellular bacteria are removed by washing with PBS and by adding 10 µg/ml gentamicin to the medium. At various times after the infection, the macrophages are lysed with 0.5% sodium deoxycholate, and suitable dilutions of the cell lysate are plated out on solmed. After incubation, the numbers of colonies are determined and the number of bacteria per ml of culture is calculated. Whereas the wild-type strain reproduces markedly in the macrophages over a period of from 7 to 24 hours, a stable *phoP* mutant is highly attenuated. A stable *phoN* mutant and a pSRM4-*phoN* insertion mutant behave like the wild type.

A pSRM4-*phoP* insertion mutant behaves like the stable *phoP* mutant. The reproduction rates are compiled in the following table

Table III:

Stability and attenuation of pSRM insertion mutants with selection in macrophages

Strain	Reproduction (x-fold) after hours			
	3	**5**	**7**	**24**
14028 Wild-type	0.1	1.0	18.8	34.7
14028 *phoN*	0.1	1.3	3.9	39.9
14028 pSRM4-*phoN*	1.6	4.2	11.3	20.8
14028 *phoP*	0.1	0.2	0.1	0.6
14028 pSRM4-*phoP*	0.1	0.1	0.1	1.5

pSRM4 is suitable for constructing SRM rPCR fragment libraries.

The rPCR (random PCR) fragment libraries are produced in the following way:

1.) Amplification of randomly generated ~500 bp fragments

After random annealing of random primer OL-30N8 [AAGTCGACGGATCCGGTACCTNNNNNNNN; N = A, T, G, or C; underlined: *Kpn*I restriction cleavage site] onto sheared chromosomal DNA from *S. typhimurium* ATCC14028 (400 ng), the primers are extended using Klenow polymerase (see fig. 11); the DNA is purified and subjected to a PCR with the primer OL-30 [AAGTCGACGGATCCGGTACCT]. Fragments of about 500 bp are eluted from an agarose gel, purified and employed for renewed amplification with the primer OL-30. This step is repeated. The exact protocol is to be found in tab. IV; the principle of the method is illustrated in fig. 11.

2.) Ligation and construction of the fragment library

Cleavage of the eluted genomic fragments with *Kpn*I and ligation with dephosphorylated, *Kpn*I-restricted pSRM4. Transformation in *E. coli* EC101 or *S. typhimurium* ATCC14028. Selection of the transformants on solmed with Tet at 30°C.

3.) Test of the fragment library by PCR and blue/white selection

Determination of the background of religated vector by means of control ligations and by means of blue/white selection on solminimal medium with BCiP (40 µg/ml). PCR of cloned inserts with primer lacZ1 and SRMgb4. The background of religated vector is about 10-25% after PCR detection and evaluation of the blue/white selection.

Table IV:

Reaction conditions for constructing a random fragment library by rPCR.

Table IV: Reaction conditions for constructing a random fragment library by rPCR.

Klenow reaction	Concentration	Final concentration	Amount [μl]
Primer (oligo 30N8)	10 μM = 10 pmol/μl	0.4 μM	2
Klenow buffer	10 ×	1 ×	2
Klenow	2.5 U/μl		0.5
dNTPs	1.25 mM	200 μM	3.5
Template: sheared chromosomal DNA		400 ng/mixture	
H_2O			ad 20

- Pipette everything except Klenow together
- 94°, 1' for denaturation, then on ice to avoid rehybridization
- 25°C, 5' for annealing
- Addition of Klenow
- Incubation 37°C, 2'
- Purification from 1.3% agarose gel with GFX gel elution kit
- Elution of the DNA in 20 + 10 μl of H_2O; DNA concentration determination

1st PCR	Concentration	Final concentration	Amount [μl]
Primer (oligo 30)	10 μM	0.8 μM	4
Buffer	10 ×	1 ×	5
dNTPs	1.25 mM	200 μM	8
Template: size-fractionated DNA from Klenow reaction			10
Taq polymerase	5 U/μl		0.4
H_2O			ad 50

- Addition of the primer only after the first PCR cycle
- 1st PCR: 95°C, 1'; 50°C, 1'; 72°C, 2'; 1 ×
- Addition of the primer
- 2nd PCR: 94°C, 40"; 64°C, 45"; 72°C, 1'; 10 ×
  3rd PCR: 94°C, 35"; 60°C, 45"; 72°C, 1'; 20 ×
- Purification from 1.3% agarose gel with GFX gel elution kit
- Elution of the DNA in 20 + 10 μl of H_2O; DNA concentration determination

2nd PCR	Concentration	Final concentration	Amount [μl]
Primer (OL30)	10 μM = 10 pmol/μl	0.8 μM	4
10 x Buffer		1 x	5
$MgCl_2$		3 mM	5
dNTPs	1.25 mM	200 μM	8
Template: size-fractionated DNA from 1st PCR	10-50 ng		
Taq polymerase	5 U/μl		0.3
H_20			ad 50
- PCR reaction: 94°C, 5'; 94°C, 35"; 60°C, 45"; 72°C, 1'; 25 x. - Loading onto 1.3% agarose gel and elution of the region around 550 bp.			

SRM fragment libraries are used to construct SRM mutant libraries.

SRM fragment libraries which have been cloned into *E.coli* EC101 are transformed into *S. typhimurium* ATCC14028. The transformants are selected with Tet on solmed at 30°C. The resulting colonies are rinsed off the plates in liqmed with Tet. To reduce the quantity of cells with free plasmid, the bacterial suspension is incubated at 37°C for 3 hours, and then plated out on solmed with Tet. Thousands of genomic insertion mutants are obtained in this way and, in their totality, represent the SRM mutant library. In order to reduce the redundancy of the SRM mutant library, independent mutant libraries are generated starting from several independent constructions of the SRM fragment library. In each case 96 SRM mutants are combined in one pool. These 96 mutants of an SRM pool can also be preserved as single clones by deep freezing.

Negative selection of SRM mutant libraries.

The SRM mutant libraries from *S. typhimurium* is selected in J774A.1 macrophages or in BALB/c mice (it is also possible to use other cell lines and other mouse strains). This entails 1 pool (96 mutants) being selected in each case. After the selection, the bacteria are initially cultivated on solmed with Tet at 37°C. Only those clones which contain genomically integrated pSRM plasmid are multiplied by this step.

Recovery of SRM fragments from genomically integrated SRM mutant libraries.

Excision of the integrated SRM plasmids from the genome takes place by incubation in liqmed with Tet at 30°C. Plasmids obtained from a selected pool of SRM mutants are referred to as drivers. Plasmids derived from the identical but <u>not</u> selected pool afford the tester. The plasmid DNA is obtained for testers and drivers by plasmid preparation from the bacteria.

Single DNA fragments cloned into pSRM4 can be identified by genetic subtraction after amplification of tester and driver fragment libraries by PCR.

The SRM fragments of the selected clones from a pool of 96 are amplified for the tester group by PCR with the primer pair SRMgb2 / SRMgb4. For the driver group, these fragments are amplified with the same primers, but the primers for amplification of the driver DNA are biotinylated at the 5' end and on the 2nd nucleotide after the 5' end. The border sequences of the tester DNA fragments can be eliminated by a subsequent restriction with *Kpn*I, so that hybridizations via the homologous primer sequences is precluded.

For the subtraction, biotinylated driver DNA and tester DNA are mixed in ratios of from 10:1 to 1000:1. The DNA in this mixture is precipitated, dissolved in 10 Fl of hybridization buffer (0.05 M HEPES pH7.5, 1mM EDTA, differing NaCl concentration), and incubated initially at 95°C for 3 minutes, followed by incubation at 65°C for 18 hours. The biotinylated homo- and heterodimers formed during this hybridization are extracted from the mixture via streptavidin-coupled magnetic particles. In order to increase the efficiency of the extraction of homologous driver DNA fragments from the tester DNA (enrichment of the tester DNA fragments to be identified), the subtraction is repeated two to three times.

For isolation and subsequent identification of the enriched tester DNA fragments, they are cloned via the *Kpn*I cleavage site into the pSRM4 vector.

Identification of SRM fragments by Southern (dot) blot.

As an alternative to the genetic subtraction it is possible to use a blotting method in order to identify SRM fragments of essential genes. For this purpose, the SRM fragments from in each case 96 SRM mutants with free pSRM plasmids are amplified by PCR (primer SRMgb2 / SRMgb4) and fixed singly on a membrane filter. For the hybridization, (1) PCR-amplified SRM fragments from the driver and (2) PCR-amplified SRM fragments from the tester are used as probe. DNA

fragments which afford a signal with the probe from the driver but not with the probe from the tester are the fragments of essential genes which are sought.

Subtractive recombination mutagenesis in *Helicobacter*

The SRM method can be used in all organisms in which pSRM4 shows temperature-sensitive replication. The SRM method can additionally be adapted for other organisms such as, for example, *Helicobacter* spp., *Campylobacter* spp., or others, by constructing SRM vectors for these organisms. Adaptation of the SRM vector for example for *Helicobacter* spp. starts from a plasmid which replicates in *Helicobacter* spp., for example pHel (Heuermann and Haas, 1995; Heuermann and Haas, 1998). The plasmid replicates like pSRM by the rolling circle mechanism, and the RepA proteins of both proteins are homologous (del Solar, *et al.*, 1993; Kleanthous, *et al.*, 1991). A temperature-sensitive variant of the plasmid is produced by the method described by Maguin et al. (Maguin, *et al.*, 1992). For this purpose, the plasmid is mutagenized *in vitro* with hydroxylamine by standard methods and transformed into *H. pylori* at the permissive temperature. Alternatively, only the *repA* gene is mutagenized, and the mutagenized DNA is then cloned into a *repA*-free plasmid. Replica plating of *H. pylori* clones containing mutangenized plasmid or *repA*, and selection of the clones by means of a plasmid-encoded antibiotic resistance leads to identification of variants of the plasmid or *repA* gene which allow replication at 30°C but not at 37°C. The mutation(s) which confer temperature sensitivity are identified by sequencing the plasmid (or the *repA* gene). The temperature sensitivity of the plasmid is additionally characterized as described above for pSRM4. The plasmid is subsequently employed like pSRM4 as vector for subtractive recombination mutagenesis in *Helicobacter* spp.

Exampl 2

Identification of essential genes by producing gene-deficient microorganisms by conditional antisense inhibition (CAI)

The CAI method can be carried out with various vectors which can replicate in the organism to be investigated and can be selected. The procedure for the method in *Salmonella typhimurium* is described by way of example below. The vector used is pCAI. pCAI contains the replication and resistance functions of pWSK29 (Wang and Kushner, 1991), or of pBluescript, or of pUC, or of other plasmids. The particular vector contains a regulatable CAI promoter and its *cis*- and *trans*-regulatory components. Transcription from this promoter is controlled by exogenous factors (inducers). The regulated CAI promoter is the *tet* promoter which is activated by the *tet* operator and the Tet repressor only in the presence of tetracycline in the growth medium (Hillen and Berens, 1994). An alternative is to use the *ara* promoter which is controlled by the *ara* operator and AraC, an activator protein, and is activated only in the presence of arabinose in the growth medium (Guzman, *et al.*, 1995). It is also possible in principle to use any other promoter regulatable by exogenous signals. Several terminator structures (Wilson and von Hippel, 1995; Yarnell and Roberts, 1999) are inserted upstream from the regulatory sequence of the CAI promoter and suppress the transcriptional activity of sequences upstream from the promoter in the absence of the inducer. One or more RNA-stabilizing elements (RSE) can be inserted downstream from the CAI promoter (Carrier and Keasling, 1997a; Carrier and Keasling, 1997b; Carrier and Keasling, 1999). The RSE are cloned in such a way that fusions with the RSE result after insertion of CAI fragments at the 5' and/or at the 3' end of the transcribed RNA. A *Kpn*I and other restriction cleavage sites (mcs) are inserted between the promoter and the 5'-RSE sequence on the one hand and the 3'-RSE sequence on the other hand and are used later for inserting CAI fragments.

A library with genomic fragments of the organism to be investigated is constructed in the CAI vector. For this purpose, genomic DNA fragments prepared by the rPCR method described in example 1 are cloned into the mcs.

The CAI gene library obtained in this way is transformed into the organism to be investigated, and individual clones are selected via the plasmid-encoded resistance marker.

The CIA method can be used with various selection and/or screening methods which are described below. The functionality of the individual screening methods is demonstrated in each case with a particular gene. The *aroA* gene, the *phoP* gene, and the *phoN* gene from *S. typhimurium* are used for this purpose. Initially, fragment libraries of (1) the *aroA* gene, (2) the *phoN* gene, and (3) the *phoP* gene in pCAI are constructed. The fragment libraries are constructed by amplifying the genomic regions of these genes, including adjacent sequences of about 500 bp on each side of the gene, by PCR. The PCR products obtained in this way are used as template for generating rPCR fragments by the method described above. The rPCR fragments are cloned, without fractionating DNA fragments of different lengths or in various groups with different lengths (fractionated by gel elution), into the *Kpn*I cleavage site in the mcs of pCAI. The fragment libraries obtained in this way are transformed into *S. typhimurium* and selected for the pCAI-encoded resistance marker. The three libraries are then tested in the following way:

(1) aroA fragment library - identification of clones by lack of growth on nutrient medium:

The clones of the *aroA* fragment library are plated onto plates with solid medium (M9 minimal medium with casamino acids [CAA, amino acid mixture) without inducer). Functioning of the *aroA* gene is necessary for the bacteria to be able to reproduce on M9 without CAA. The resulting colonies are transferred by replica plating to M9 with CAA and to M9 without CAA, in each case with inducer. Clones in which fragments of the *aroA* fragment library are transcribed in antisense orientation from the CAI promoter and synthesize RNA which inhibit the translation of the genomic *aroA* gene reproduce on M9 with CAA with inducer but not on M9 without CAA with inducer. Clones which do not reproduce on M9 without CAA are isolated and analyzed further.

(2) phoN fragment library - identification of clones by altered biochemical (color) reaction on solid medium:

The clones of the *phoN* fragment library are transferred by replica plating to M9 medium with BCiP and with inducer, and to M9 with BCiP without inducer, The PhoN activity converts BCiP into a blue dye. Clones in which translation of the *phoN* gene are inhibited by asRNA remain whitish on M9 medium with BCiP with inducer, whereas they develop a blue color on M9 medium with BCiP without inducer.

(3) phoP fragment library - identification of clones from negatively selected pools:

Functioning of the *phoP* gene is necessary for the bacteria to be able to reproduce in macrophages. The clones of the phoP fragment library are combined in one pool. The pool is divided into two identical groups which are passaged into J774A.1 macrophages. For one group, inducer is added to the culture medium of the macrophages during the infection. The second group is passaged in J774A.1 macrophages without added inducer. (see example 1. for description of the infection experiment). After the infection (lasting between 7 and 24 hours), the pCAI plasmids are prepared from both mixtures. The CAI fragments of both mixtures are amplified by PCR with primers which hybridize adjacent to the cloned fragments on the pCAI vector, and are then separated from one another by genetic subtraction (see example1). Fragments leading to inactivation of *phoP* are isolated in this way.

(4) Application of tests 1-3 to a genomic fragment library:

The CAI method is then applied to a complete genomic fragment library in *S. typhimurium.* The clones with the CAI fragment library are tested by the three methods (1-3) described above.

Example 3

Enrichment of *H. pylori* genes coding for secreted/exported polypeptides

Production of an H. pylori gene library in the minimal vector pMin2:

The *H. pylori* wild-type strain 69A is used as initial strain. The bacteria are grown on serum plates (see Westblom et al., 1991) at a temperature of 37°C in an atmosphere of 5% O_2, 10% CO_2, 85% N_2. The chromosomal DNA is isolated by the method of Leying et al. (1992), with the DNA subsequently being purified on a cesium chloride gradient. 50 µg of the purified chromosomal DNA is partially cleaved with the restriction endonucleases Sau3A and HpaII, the DNA fragments are fractionated in an agarose gel, and the fragments with a size of from 3 to 6 kbp are eluted from the gel using the Geneclean II kit (Bio101). The isolated DNA fragments are cloned into the BglII- and ClaI-cut pMin2 vector(Kahrs et al., 1995) and transformed by electroporation into the E. coli strain E181 into which the plasmid pTnMax9 had previously been transferred. In total, about 4 000 clones are generated in such an approach. The E. coli strain E181 is a derivative of the strain HB101 (Boyer and Roulland-Dussoix, 1969) and contains the lysogenic λ phage λCH616 for replication of the pTnMax9 plasmid.

Serum plate composition

Suspend 36 g of GC agar (base) in 910 ml of distilled water and autoclave

(-) allow to cool to about 45°C

(-) add 10 ml of vitaminmix

(-) add in each case 1 ml of the antibiotic stock solution

  (-) vancomycin (10 mg/l in double-distilled water)

  (-) nystatin (0.793 mg/l in DMF)

  (-) trimethoprim (5 mg/l in DMF)

  (-) amphotericin (5mg/l in DMF)

  *DMF=dimethylformamide

(-) add 90 ml of serum or

(-) add 8% horse blood = 80 ml

(-) or an equal quantity of human blood

- *Preparation of the components:*

- Antibiotic stock solution

Vancomycin	100 mg in 10 ml of double-distilled water	
Nystatin	7.93 mg in 10 ml of DMF	add to each 1 ml per liter GC
Trimethoprim	50 mg in 10 ml of DMF	agar
Amphotericin	50 mg in of 10 ml DMF	

(Storage in a refrigerator for a max. of 8 weeks)

Vitamin mix (concentrate)

Dextrose (D-glucose)	100 g	
L-Glutamine	10 g	
Cysteine HCl ($C_3H_7NO_2S$ x HCl x H_2O)	26 g	
Cocarboxylase	100 mg	dissolve in 50 ml of double-distilled
Fe $(NO_3)_3$	20 mg	water
Thiamine HCl	3 mg	
DPN NAD	250 mg	
Vitamin B_{12}	10 mg	
L-Cysteine ($C_6H_{12}N_2O_4S_2$)	1.1 g	
Adenine	1.0 g	dissolve in 15 ml HCl
Guanine Cl	30 mg	(32% strength)
Uracil	500 mg	
L-Arginine HCl	150 mg (sterilize by filtration, divide into 10 ml	
p-Aminobenzoic acid	13 mg portions + refrigerate)	

Genetic enrichment of secreted/excreted H. pylori gene products:

The transposon Tn*Max9* located on pTn*Max9* is equipped with the genetic marker β-lactamase. This marker is situated on the transposon in such a way that selection of successful transposon insertion is made possible thereby when this transposon is present in the correct reading frame of genes whose products are secreted or

exported by the *E. coli* strain E145^{rif}. Insertion of the transposon into such a gene leads to a gene fusion between the target gene and the marker, the fusion protein being expelled from the cell by a secretion and/or export signal determined in the target gene, and the activity of the integral reporter gene being displayed. In the case of β-lactamase, the clones can be detected directly through the development of a resistance to ampicillin. The Tn*Max9* transposon on pTn*Max9* is activated by IPTG.

Transposon mutagenesis of the gene library is carried out in pools of up to 20 individual clones. The respective pools are plated out on LB plates to which 100 μM IPTG, 15 μg/ml chloramphenicol and 15 μg/ml tetracycline are added. In a second step, the Tn*Max9* mutagenized pMin2 plasmids are transferred by conjugation into the E. coli strain E145^{rif} because the latter are equipped with an appropriate mob signal (*oriT*). The pTn*Max9* plasmids are, by contrast, not transferred. Consequently, specific multiplication of the mutagenized pMin2 gene library takes place in *E. coli* E145^{rif}. The corresponding transconjugants are selected on LB medium with 15μg/ml chloramphenicol, 15μg/ml tetracycline and 100 μg/ml rifampicin. In total, 500 - 1 000 transconjugants are combined in 2 ml of LB medium and grown in appropriate dilutions (10^{-1} -10^{-2}) on LB plates to which 50 μg/ml ampicillin are added. After the plates have been cultivated at 37°C for 36 hours, 200-300 ampicillin-resistant clones with transposon-inserted plasmids are obtained in this complete batch.

Production of gene-deficient *H. pylori* which are mutated in genes encoding secreted/excreted polypeptides, and identification of those genes having an essential biological function.

Production of gene-deficient H. pylori mutants:

Gene-specific mutants can be generated by introducing the isolated plasmids with the Tn*Max9*-mutated *H. pylori* genes coding for secreted/excreted polypeptides

into a wild-type *H. pylori* strain. Owing to the cloned *H. pylori* gene sequences on the plasmids, there is genomic insertion of the Tn*Max9* transposon into the chromosomal target gene, and thus inactivation thereof, if there is a double homologous recombination event. The genetic marker on the transposon makes it possible to select this process because the pMin2 plasmid is not replicated in *H. pylori*.

In the procedure, the wild-type *H. pylori* strain 69A is transformed in individual mixtures with the isolated individual plasmids, making use of the natural competence of the bacterium to take up DNA (Haas et al., 1993). Starting from a preculture on serum plates according to the standard cultivation conditions, the bacteria are taken up in BHI medium and grown to an optical density at 550 nm of 0.1 at 37°C under microaerophilic conditions. 100 to 500 ng of purified plasmid DNA are added to each of the individual culture mixtures, and the culture is continued overnight.

Characterization of the biological function of the gene-deficient H. pylori mutants in the growth test:

After the cultivation, the individual mixtures are plated out on serum plates to which 4 µg/ml chloramphenicol are added. In relation to the growth properties of the individual mutants compared with the wild-type strain, 3 categories can be distinguished:(1) mutants which do not grow; (2) mutants which form smaller colonies;(3) mutants which develop colonies of normal size. These results can be obtained reproducibly for any initial plasmid. For unambiguous assessment of the biological significance of the gene-deficient mutants of category 1 they are subjected to the CAI method. The gene-deficient mutants of category 2 and 3 are analyzed in the other biological test systems.

Example 4

Establishment of the identity of the *H. pylori* genes coding for secreted/exported polypeptides.

The primary structure of the identified *H. pylori* genes is established by using the respective initial plasmids from the *E. coli* strain. The plasmids are isolated from these strains, and the nucleotide sequence of the target genes is determined by sequencing the regions above and below the insertion site of the transposon in the target gene. The reading frame of the target gene can be established directly because the transposon-encoded ss-lactamase gene undergoes active fusion with the gene product of the target gene(see above). The sequencing is carried out using an ABI automatic sequencer in accordance with the manufacturer's information and using the following sequence primers: M13-F (GTAAAACGACGGCCAGT) and M13-RP1 (CAGGAAACAGCTATGACC). For further characterization of the genes, use is made of the database Genebank of the GCG program, for example for identifying known homologous genes of other microorganisms (FASTA), identifying potential signal peptide regions (SPSCAN) or for identifying lipoproteins (MOTIFS).

It was not possible to establish the complete gene sequence for some of the identified clones because the cloned DNA fragment did not contain the whole gene. The available DNA fragment is employed to isolate a DNA fragment with the complete gene from the original gene library. It is then possible for the missing gene sequences from these clones to be amplified with a gene-specific primer and a vector-specific primer and then be directly sequenced. For this purpose, chromosomal DNA is isolated from the *H. pylori* strain 69A, 25-35 µg thereof are partially digested with Bsp143I (Sau3AI isoschizomer), and resulting fragments with a size of from 2 to 8 kbp are isolated. For the cloning, the vector pACYC184 is cut with BamHI, dephosphorylated with shrimp alkaline phosphatase and then ligated with the isolated fragments. After transformation into the *E. coli* strain 0466 and selection on 30 µg/ml chloramphenicol, 5560 individual colonies are isolated and analyzed by the polymerase chain reaction.

In this case, primers for a product size of less than/equal to 700 bp are derived from the known gene fragments as close as possible to the missing piece of gene. These primers are used to produce amplification products which contain a clone which comprises both primer binding sites. The plasmids are then isolated from the positive clones and are verified either by sequencing or by dot blot analysis.

Sequencing

The technique of primer walking is used to complete the DNA sequence of genes which cannot be determined completely after the sequencing reactions with the universal primers M13F and M13RP1. For this purpose, an oligonucleotide which binds about 50 to 200 bp upstream from the 3' end of the known sequence section is deduced and employed as primer for another sequencing reaction according to the dye terminator principle (Amersham Pharmacia Biotech). For the example of the completely determined DNA sequence of the gene *hpc052* of *H. pylori* 69A, this is achieved by the oligonucleotide primers Mu052a (5'-AGGCTAAAGACGTGTTAG-3') and Mu052b (5'-CTAGCGTGGAATTAGCC-3').

Expression cloning

The gene products of the identified gene sequences are characterized by heterologous expression of the gene sequences in attenuated salmonella vaccine strains. The polymerase promoter system of bacteriophage T7 (Tabor & Richardson 1985) is used for this purpose. The gene to be expressed is amplified from chromosomal DNA by a polymerase chain reaction(PCR) with specific oligonucleotide primers. Selection of the gene-specific oligonucleotide primers depends on the presence of a signal peptide for the Sec-dependent translocation of the gene product into the periplasm. If one of these is present, the oligonucleotide primer pair is chosen so that there is amplification of a fragment of the gene to be expressed which codes for the mature protein without signal peptide. If no signal sequence is present, the complete open reading frame is amplified.

The amplified gene fragments are inserted into a plasmid vector derived from pBR322. The insertion takes place in such a way that the cloned gene fragments are each present as fusion with a sequence section which codes for the detection epitope RGSHis^4 under the transcriptional control of the bacteriophage T7 promoter. The expression of the gene products is investigated in the *Salmonella typhimurium* vaccine strain SL3261::pYZ84 (Yan & Meyer 1996) which is transformed with the expression vectors produced in the above manner. The RGSHis^4 fusion proteins are detected by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent Western blot analysis with a monoclonal antibody (RGS.His) which is specific for the RGSHis^4 epitope.

An expression cloning with subsequent Western blot analysis is described by way of example for the gene sequence of HPC001. For this purpose, the gene fragment of the *hpc001* gene which codes for the mature HPC001 protein (without signal peptide) is amplifed by PCR using the oligonucleotide primer LAT50 (5'-GATCAGATCTACATATGTTTATCATTCCCTCTCGC-3') and LAT51 (5'-GATCGGTACCAAAACCTTAATGCGTTGCT-3'). The amplicon 0.9 kb in size is hydrolyzed with *Bgl*II and *Acc*65I and inserted into the vector pLAT289 which has been hydrolyzed with *Bam*HI and *Acc*65I. The resulting expression vector pMSC34 codes for a fusion protein which consists of the *hpc001* gene lacking the signal sequence of HPC001 and of the RGSHis^4 epitope. The expression product has a molecular mass of 34 kDa.

Example 5

"Coccoid helicobacter" biological model

Production of coccoid forms.

Experimental conversion of the spiral-shaped helicobacter form into a coccoid form can take place in various ways. In particular, the transformation can be observed under stress conditions, for example in the presence of sublethal doses of an effective antibiotic, if there is substrate deficiency or in a normal air atmosphere.

The transformation process can thus be understood as a protective reaction of the bacteria. These forms can in fact also be observed in the tissue of infected people. This possibly comprises a resistant form. Loss or delayed or incomplete expression of this property, for example through loss or inactivation of a gene or gene product, may have fatal disadvantages for the bacterium in the fight for survival.

Experimental transformation of the spiral-shaped helicobacter into a coccoid form is usually started from a fresh plate culture which has been grown under ideal conditions (serum plates, 37°C, microaerophilic atmosphere (5% O_2, 10% CO_2, 85% N_2)) and thus contains predominantly spiral-shaped helicobacter. In the following example, the harvested bacteria are transferred into a minimal medium which is composed of water, for example sterilized tap water, and x% (v/v) of a complex protein solution, for example fetal calf serum(FCS). Depending on the FCS concentration of x, the bacteria are transformed within a few hours or in several days completely into coccoid forms. Usual concentrations are 10 – 30% (v/v) FCS in water. In the usual arrangement, the bacteria are cultivated in a defined number, for example at an OD_{550} of 0.2, in 5 ml of minimal medium at 37°C in a cell culture cabinet, adjusting the proportion of CO_2 in the air to 10%. The progress of the development process is documented on the one hand by microscopy and by determinations of growth. In phase contrast microscopy, the morphology (coccoid _ spiral-shaped) and the motility (coccoid + non-motile; spiral-shaped + motile; spiral-shaped + non-motile) is determined proportionately for the bacteria in the culture mixture, for example in a counting chamber. Where appropriate, specific stains are carried out in order also to detect indirectly structural changes, for example in the bacterial envelope and determine the number of dead microbes. Since coccoid forms do not grow on conventional media, for example serum plates, it is possible to establish proportionately at different times the proportion of vital and still reactivatable forms (colony forming units; CFU).

R activation of coccoid forms.

The coccoid bacteria are activated in a medium which is described in EP 0900839. The bacteria are plated out in a defined number, for example 500 – 1 000 bacteria, on this medium and incubated under microaerophilic conditions at 37°C for several days. The development process is again documented by microscopy (see above) and by determination of the CFU in this approach.

Finally, these investigations can also be carried out in an animal model. It is to be assumed that the reactivation and selection mechanisms operating in the animal differ from those in the cultivation model described above. For example, other or additional genes or gene groups may be involved in the process. The produced coccoid forms are administered orally in a defined number, for example 10^8 bacteria, in an aqueous solution to a mouse, for example Balb/c, for example using a stomach tube. After a defined time, for example after 2 - 4 weeks, the infected animals are sacrificed, and the stomach is removed, freed of contents and investigated for colonization histologically and by growth tests (see above). It is possible in the histological investigation to determine the proportion of coccoid and spiral-shaped bacteria. The experiment ought preferably to be carried out with helicobacter strains of type I, because the type II strains are less virulent.

Identification of essential helicobacter genes.

The described biological systems (A+B) can be employed, for example, for analyzing defined helicobacter mutants produced by the methods described. The investigation takes place in each case by comparision with the respective wild-type strain. If, for example, a mutant is unable to transform into coccoid forms or if there is no reactivation of the coccoid form or if the transformations occur with long time delays, a facultatively essential function can be inferred for the mutated gene. In the reactivation of coccoid forms in an animal model it is possible to employ several mutants at the same time, for example 10 strains, per experiment. The precondition

is that recognition of this strain in the tissue is ensured by suitable methods, for example by appropriate PCR methods or by *in situ* hybridization.

References

Agterberg, M., Adriaanse, H., van Bruggen, A., Karperien, M., Tammassen, J. (1990). Outer-membrane PhoE protein of Escheicia coli K-12 as an exposure vector: possibilities and limitations. Gene 88: 37-45.

Akerley, B.J., Rubin, E.J., Camilli, A., Lampe, D.J., Robertson, H.M. and Mekalanos, J.J. (1998). Systematic identification of essential genes by in vitro mariner mutagenesis. Proc Natl Acad Sci U S A, 95:8927-32

Arigoni, F., Talabot, F., Peitsch, M., Edgerton, M.D., Meldrum, E., Allet, E., Fish, R., Jamotte, T., Curchod, M.L. and Loferer, H. (1998). A genome-based approach for the identification of essential bacterial genes. Nat Biotechnol, 16: 851-856.

Barry, E.M., Gomez-Duarte, O., Chatfield, S., Rappuoli, R., Pizza, M., Losonsky, G., Galen, J., Levine, M.M. (1996). Expression and immunogenicity of pertussis toxin S1 subunit-tetanus toxin fragment C fusions in Salmonella typhi vaccine strain CVD908. Infect. Immun. 64: 4172-81.

Bergler, H., Fuchsbichler, S., Hogenauer, G., Turnowsky, F. (1996). The enoyl-[acyl-carrier-protein] reductase (FabI) of Escherichia coli, which catalyzes a key regulator step in fatty acid biosynthesis, acccepts NADH and NADPH as cofactors and is inhibited by palmitoyl-CoA. Eur. J. Biochem. 242: 689-694.

Berry, V., Jennings, K. and Woodnutt, G. (1995) Bactericidal and morphological effects of amoxicillin on Helicobacter pylori. Antimicrob Agents Chemother 39: 1859-1861.

Bode, G., Mauch, F. and Malfertheiner, P. (1993) The coccoid forms of Helicobacter pylori. Criteria for their viability. Epidemiol Infect 111: 483-490.

Cao, J., Li, Z.Q., Borch, K., Petersson, F. and Mardh, S. (1997) Detection of spiral and coccoid forms of Helicobacter pylori using a murine monoclonal antibody. Clin Chim Acta 267: 183-196.

Carrier, T.A. and Keasling, J.D. (1997a). Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. Biotechnol. Prog. 13: 699-708.

Carrier, T.A. and Keasling, J.D. (1997b). Engineering mRNA stability in *E. coli* by the addition of synthetic hairpins using a 5' cassette system. Biotechnol. Bioeng. 55: 577-580.

Carrier, T.A. and Keasling, J.D. (1999). Library of synthetic 5' secondary structures to manipulate mRNA stability in *Escherichia coli.* Biotechnol. Prog. 15: 58-64.

Cellini, L., Allocati, N., Angelucci, D., Iezzi, T., Di Campli, E., Marzio, L. and Dainelli, B. (1994) Coccoid Helicobacter pylori not culturable in vitro reverts in mice. Microbiol Immunol 38: 843-850.

Chan, W.Y., Hui, P.K., Leung, K.M., Chow, J., Kwok, F. and Ng, C.S. (1994) Coccoid forms of Helicobacter pylori in the human stomach. Am J Clin Pathol 102: 503-507.

Chow, W.Y. and Berg, D.E. (1988). Tn5tac1, a derivative of transposon Tn5 that generates conditional mutations. Proc Natl Acad Sci U S A, 85:6468-72

Cirillo, J.D., Stover, C.K., Bloom, B.R., Jacobs, W.R. Jr., Barletta, R.G. (1995). Bacterial vaccine vectors and bacillus Calmette-Guerin. Clin. Infect. Dis. 20: 1001-1009.

Dalbey, R.E (1991). Leader peptidase. Mol. Microbiol. 5: 2855-2860.

Darji, A., Guzmán, C.A., Gerstel, B., Wachholz, P., Timmis, K.N., Wehland, J., Chakraborty, T., Weiss, S. (1997). Oral somatic transgene vaccination using attenuated S. typhimurium. Cell 91: 765-775.

Das, A., Court, D. and Adhya, S. (1976). Isolation and characterization of conditional lethal mutants of Escherichia coli defective in transcription termination factor rho. Proc Natl Acad Sci U S A, 73:1959-63

Davenport MP, Ho Shon IA, Hill AV (1995). An empirical method for the prediction of T-cell epitopes. Immunogenetics 1995;42(5):392-7

del Solar, G., Moscoso, M. and Espinosa, M. (1993). Rolling circle-replicating plasmids from gram-positive and gram-negative bacteria: a wall falls. Mol. Microbiol. 8: 789-796.

Donelli, G., Matarrese, P., Fiorentini, C., Dainelli, B., Taraborelli, T., Di Campli, E., Di Bartolomeo, S. and Cellini, L. (1998) The effect of oxygen on the growth and cell morphology of Helicobacter pylori. FEMS Microbiol Lett 168: 9-15.

Dorner, F. (1995). An overview of vaccine vectors. Dev. Biol. Stand. 84: 23-32.

Drlica K., Zhao, X. (1997). DNA gyrase, topoisomerase IV, and the 4-quinolones. Microbiol. Mol. Biol. Rev. 61: 377-392.

Eaton, K.A., Catrenich, C.E., Makin, K.M. and Krakowka, S. (1995). Virulence of coccoid and bacillary forms of Helicobacter pylori in gnotobiotic piglets. JID 171: 459-462.

Ermak, T.H., Giannasca, P.J., Nichols, R., Myers, G.A., Nedrud, J., Weltzin, R., Lee, C.K., Kleanthous, H., Monath, T.P. (1998). Immunization of mice with urease vaccine affords protection against Helicobacter pylori infection in the absence of

antibodies and is mediated by MHC class II-restricted responses. J. Exp. Med. 188(2): 2277-2288.

Franciso, J.A., Earhart, C.F., Georgiou, G. (1992). Transport and anchoring of beta-lactamase to the external surface of Escherichia coli. Proc. Natl. Acad. Sci. USA 89: 2713-2717.

Gabelsberger, J., Knapp, B., Bauersachs, S., Enz, U.I., von Specht, B.U., Domdey, H. (1997). A hybrid outer membrane protein antigen for vaccination against Pseudomonas aeruginosa. Behring Inst. Mitt. 98: 302-314.

Gonzalez, C.R., Noriega, F.R., Huerta, S., Santigao, A., Vega, M., Panagua, J., Ortiz-Navarrete, V., Isibasi, A., Levine, M.M. (1998). Immunogenicity of a Salmonella typhi CVD908 candidate vaccine strain expressing the major surface protein gp63 of Leishmania mexicana mexicana. Vaccine, 16: 1043-52.

Goodwin, C.S., Armstrong, J.A., Chilvers, T., Peters, M., Collins, M.D., Sly, L., McConnell, W., Harper, W.E.S. (1989). Transfer of Campylobacter pylori and Campylobacter mustelae to Helicobacter gen. nov. as Helicobacter pylori comb. nov and Helicobacter mustelae comb. nov., respectively. Int. J. Syst. Bact. 39: 397-405.

Gupta, R.K., Siber, G.R. (1995). Adjuvants for human vaccines – current status, problems and future prospects. Vaccine 13: 1263-1276.

Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. 177: 4121-4130.

Haq, T.A., Mason, H.S., Clements, J.D., Arntzen, C.J. (1995). Oral immunization with a recombinant bacterial antigen produced in transgenic plants. Science 268: 714-716.

Harris, S.D., Cheng, J., Pugh, T.A. and Pringle, J.R. (1992). Molecular analysis of Saccharomyces cerevisiae chromosome I. On the number of genes and the identification of essential genes using temperature-sensitive-lethal mutations. J Mol Biol, 225:53-65

Hartikka, J., Sawdey, M., Cornefert-Jensen, F., Margalith, M., Barnhart, K., Nolasco, M., Vahlsing, H.L., Meek, J., Marquet, M., Hobart, P., Norman, J., Manthorpe, M. (1996). An improved plasmid DNA expression vector for direct injection into skeletal muscle. Hum. Gene Ther. 7: 1205-1217.

Heuermann, D. and Haas, R. (1995). Genetic organization of a small cryptic plasmid of *Helicobacter pylori.* Gene 165: 17-24

Heuermann, D. and Haas, R. (1998). A stable shuttle vector system for efficient genetic complementation of *Helicobacter pylori* strains by transformation and conjugation. Mol. Gen. Genet. 257: 519-528

Hiatt, A., Ma, J.K. (1993). Characterization and applications of antibodies produced in plants. Int. Rev. Immunol. 10: 139-152.

Hill, J., Leary, S.E., Griffin, K.F., Williamson, E.D., Titball, R.W. (1997). Regions of Yersinia pestis V antigen that contribute to protection against plague identified by passive and active immunization. Infect. Immun. 65: 4476-4482.

Hillen, W. and Berens, C. (1994). Mechanisms underlying expression of Tn10 encoded tetracycline resistance. Annu. Rev. Microbiol. 48: 345-369.

Hou, Y., Lin, Y.P., Sharer, J.D. and March, P.E. (1994). In vivo selection of conditional-lethal mutations in the gene encoding elongation factor G of Escherichia coli. J Bacteriol, 176:123-9

Hughes, E.E., Gilleland, L.B., Gilleland, H.E. Jr. (1992). Synthetic peptides representing epitopes of outer membrane protein F of Pseudomonas aeruginosa that elicit antibodies reactive with whole cells of heterologous immunotype strains of P. aeruginosa. Infect. Immun. 60: 3497-3503.

Joyce, C.M. and Grindley, N.D. (1984). Method for determining whether a gene of Escherichia coli is essential: application to the polA gene. J Bacteriol, 158:636-43

Kalinke, U., Krebber, A., Krebber, C., Bucher, E., Plückthun, A., Zinkernagel, R.M., Hengartner, H. (1996). Monovalent single-chain Fv fragments and bivalent miniantibodies bound to vesicular stomatitis virus protect against lethal infection. Eur. J. Immunol. 26: 2801-2806.

Khin, M.M., Ringner, M., Aleljung, P., Wadström, T. and Ho, B. (1996) Binding of human plasminogen and lactoferrin by Helicobacter pylori coccoid forms. J Med Microbiol 45: 433-439.

Klauser, T., Pohlner, J., Meyer, T.F. (1990). Extracellular transport of cholera toxin B subunit using Neisseria IgA protease beta-domain: conformation-dependent outer membrane translocation. EMBO J. 9: 1991-1999.

Kleanthous, H., Clayton, C.L. and Tabaqchali, S. (1991). Characterization of a plasmid from *Helicobacter pylori* encoding a replication protein common to plasmids in Gram-positive bacteria. Mol. Microbiol. 5: 2377-2389.

Kok, J., van der Vossen, J.M. and Venema, G. (1984). Construction of plasmid cloning vectors for lactic streptococci which also replicate in *Bacillus subtilis* and *Escherichia coli.* Appl. Environ. Microbiol. 48: 726-731.

Kraiss, A., Schlor, S., Reidl, J. (1998). In vivo transposon mutagenesis in Haemophilus influenza. Appl. Environ. Microbiol. 64: 4697-4702.

Kujau, M.J., Hoischen, C., Riesenberg, D., Gumpert, J. (1998). Expression and secretion of functional miniantibodies McPC603scFvDhlx in cell-wall-less L-form strains of Proteus mirabilis and Escherichia coli: a comparison of the synthesis capacitites of L-form strains with an E. coli producer strain. Appl. Microbiol. Biotechnol. 49: 51-58.

Kusters, J.G., Gerrits, M.M., Van Strijp, J.A. and Vandenbroucke-Grauls, C.M. (1997) Coccoid forms of Helicobacter pylori are the morphologic manifestation of cell death. Infect Immun 65: 3672-3679.

Lang, H., Korhonen, T.K. (1997). The OmpS maltoporin of Vibrio cholerae as carrier of foreign epitopes. Berhring Inst. Mitt. 98: 400-409.

Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G. and Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. J. Bacteriol. 177: 7011-7018.

Lee, A., Fox, J., Hazel, S. (1993). Is gastric lymphoma an infectious disease?. Hum. Pathol. 24: 569-70.

Liljeqvist, S., Stahl, S., Andreoni, C., Binz, H., Uhlen, M., Murby, M. (1997). Fusions to the cholera toxin B subunit: influence on pentamerization and GM1 binding. J. Immunol. Methods 210: 125-135.

Ling, Y.S., Guo, Y.J., Li, J.D., Yang, L.K., Luo, Y.X., Zhen, L.Q., Qiu, S.B., Zhu, G.F. (1998). Serum and egg yolk IgG antibody titers from laying chickens vaccinated with Pasteurella multocida. Avian Dis. 42: 186-189.

Ma, J.K., Lehner, T., Stabila, P., Fux, C.I., Hiatt, A. (1994). Assembly of monoclonal antibodies with IgG1 and IgA heavy chain domains in transgenic tobacco plants. Eur. J. Immunol. 24: 131-138.

Maguin, E., Duwat, P., Hege, T., Ehrlich, D. and Gruss, A. (1992). New thermosensitive plasmid for gram-positive bacteria. J. Bacteriol. 174: 5633-5638.

Malfertheiner, P. (1994). Helicobacter pylori – Von der Grundlage zur Therapie. Georg Thieme Verlag, Stuttgart, New York; p. 1-104.

Manoil, C., Mekalanos, J.J., Beckwith J. (1990). Alkaline phosphatase fusions: sensors of subcellular location. J. Bacteriol. 172: 515-518.

Marshall, B.J., Royce, H., Annear, D.I., Goodwin, C.S., Pearman, J.W., Warren, J.R., Armstrond, J.A. (1984). Original isolation of Campylobacter pyloridis from human gastric mucosa. Microbios Lett. 25: 82-88.

Mason H.S., Arntzen, C.J. (1995). Transgenic plants as vaccine production systems. Trends Biotechnol. 13: 388-392.

Mekalanos, J.J. (1994). Live attenuated vaccine vectors. Int. J. Technol. Assess Health Care 1994 Winter 10: 131-142.

Murphy, C.K., Stewart, E.J. and Beckwith, J. (1995). A double counter-selection system for the study of null alleles of essential genes in Escherichia coli. Gene, 155:1-7

Narikawa, S., Kawai, S., Aoshima, H., Kawamata, O., Kawaguchi, R., Hikiji, K., Kato, M., Iino, S. and Mizushima, Y. (1997) Comparison of the nucleic acids of helical and coccoid forms of Helicobacter pylori. Clin Diagn Lab Immunol 4: 285-290.

Nomura, A., Stemmermann, G.N., Chyou, P.H., Kato, I., Perez-Perez, G.I., Blaser, M.J. (1991). Helicobacter pylori infection and gastric carcinoma among japanese americans in Hawaii. N. Engl. J. Med. 325: 1132-1136.

O'Hagan, D.T. (1998). Recent advances in vaccine adjuvants for systemic and mucosal administration. J. Pharm. Pharmacol. 50: 1-10.

Pack, P., Kujau, M., Schroeckh, V., Knupfer, U., Wenderoth, R., Riesenberg, D., Plückthun, A. (1993). Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of Escherichia coli. Biotechnology 11: 1271-1277.

Parsonnet, J., Friedman, G.D., Vandersetten, D.P., Chang, Y., Vogelman, H.H., Orentreich, N., Sibley, R.K. (1991). Helicobacter pylori infection and the risk of gastric carcinoma. N. Engl. J. Med. 325: 1227-1131.

Polissi, A. and Georgopoulos, C. (1996). Mutational analysis and properties of the msbA gene of Escherichia coli, coding for an essential ABC family transporter. Mol Microbiol, 20:1221-33

Pugsley, A.P. (1993). The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. 57: 50-108.

Ray, S.K., Parimoo, P., Nandi, I., Venkidesh, R., Murthy, M.V., Srisunder, S., Anand, V., Rajesh, K., Rangaraj, N. (1998). "Resultant bond moment"-as a newly developed electronic parameter in the design of antibacterial, antiprotozoal nitroimidazole derivatives. Arzneimittelforschung, 48: 286-290.

Sasse, M., Kruger, M., Schade, R., Hlinak, A. (19989). Generation and characterization of avian vitelline antibodies against lipopolysaccharide and lipid A. 1. Induction and preparation of specific egg yolk antibodies (IgY) against endotoxins. Berl. Munch Tierärztl. Wochenschrift 111: 121-126.

Segal, E.D., Falkow, S. and Tompkins, L.S. (1996) Helicobacter pylori attachment to gastric cells induces cytoskeletal rearrangements and tyrosine phosphorylation of host cell proteins. Proc Natl Acad Sci U S A 93: 1259-1264.

Solnick, J.V., Tompkins, L.S. (1993). Helicobacter pylori and gastroduodenal disease: pathogenesis and host-parasite interaction. Infect.Ag. Dis. 1: 294-309.

Sorberg, M., Nilsson, M., Hanberger, H. and Nilsson, L.E. (1996) Morphologic conversion of Helicobacter pylori from bacillary to coccoid form. Eur J Clin Microbiol Infect Dis 15: 216-219.

Tacket, C.O., Mason, H.S., Losonsky, G., Clements, J.D., Levine M.M., Arntzen, C.J. (1998). Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato. Nat. Med. 4: 607-609.

Tadayyon, M., Broome-Smith, J.K. (1992). TnblaM – atronsposon for directly tagging bacterial genes encoding cell envelope and secreted proteins. Gene 111: 21-26.

Tang, D.C., DeVit, M., Johnston, S.A. (1992). Genetic immunization is a simple method for eliciting an immune response. Nature 356: 152-154.

Van Engelen, F.A., Schouten, A., Molthoff, J.W., Roosien, J., Salinas, J., Dirkse, W.G., Schots, A., Bakker, J., Gommers, F.J., Jongsma, M.A. et al. (1994). Coordinate expression of antibody subunit genes yields high levels of functional antibodies in roots of transgenic tobacco. Plant Mol. Biol. 26: 1701-1710.

Wang, R.F. and Kushner, S.R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli.* Gene 100: 195-199.

Wang, X., Sturegard, E., Rupar, R., Nilsson, H.O., Aleljung, P.A., Carlen, B., Willen, R. and Wadström, T. (1997) Infection of BALB/c A mice by spiral and coccoid forms of Helicobacter pylori. J Med Microbiol 46: 657-663.

Warren, J.R., Marshall, B. (1983). Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet i: 1273-1275.

Wilson, K.S. and von Hippel, P.H. (1995). Transcription termination at intrinsic terminators: the role of the RNA hairpin. Proc. Natl. Acad. Sci. USA 92: 8793-8797.

Yarnell, W.S. and Roberts, J.W. (1999). Mechanism of intrinsic transcription termination and antitermination. Science 284: 611-615.